SCHOOL OF INFORMATION

Data, Models, and Knowledge in Climate Science:

Historical Lessons and Contemporary Controversies

Paul N. Edwards
School of Information and Dept. of History, University of Michigan

Introduction

- **some ideas about climate knowledge**

Weather forecasting

Computers change everything

Climate data and climate models

Data friction

Climate controversies

Conclusion

Some ideas about climate knowledge

- *A Vast Machine*
  - A history of climate science as a global knowledge infrastructure — data, theory, models, devices
    - Infrastructure: reliable, ubiquitous, invisible except on breakdown
- Thinking globally
- Making global data
- Making data global
- Friction: data and computation
- Infrastructural inversion as a method
  - …and the problems it creates

Introduction

Weather forecasting

- **some history of technologies and techniques**
- **making global data**

Computers change everything

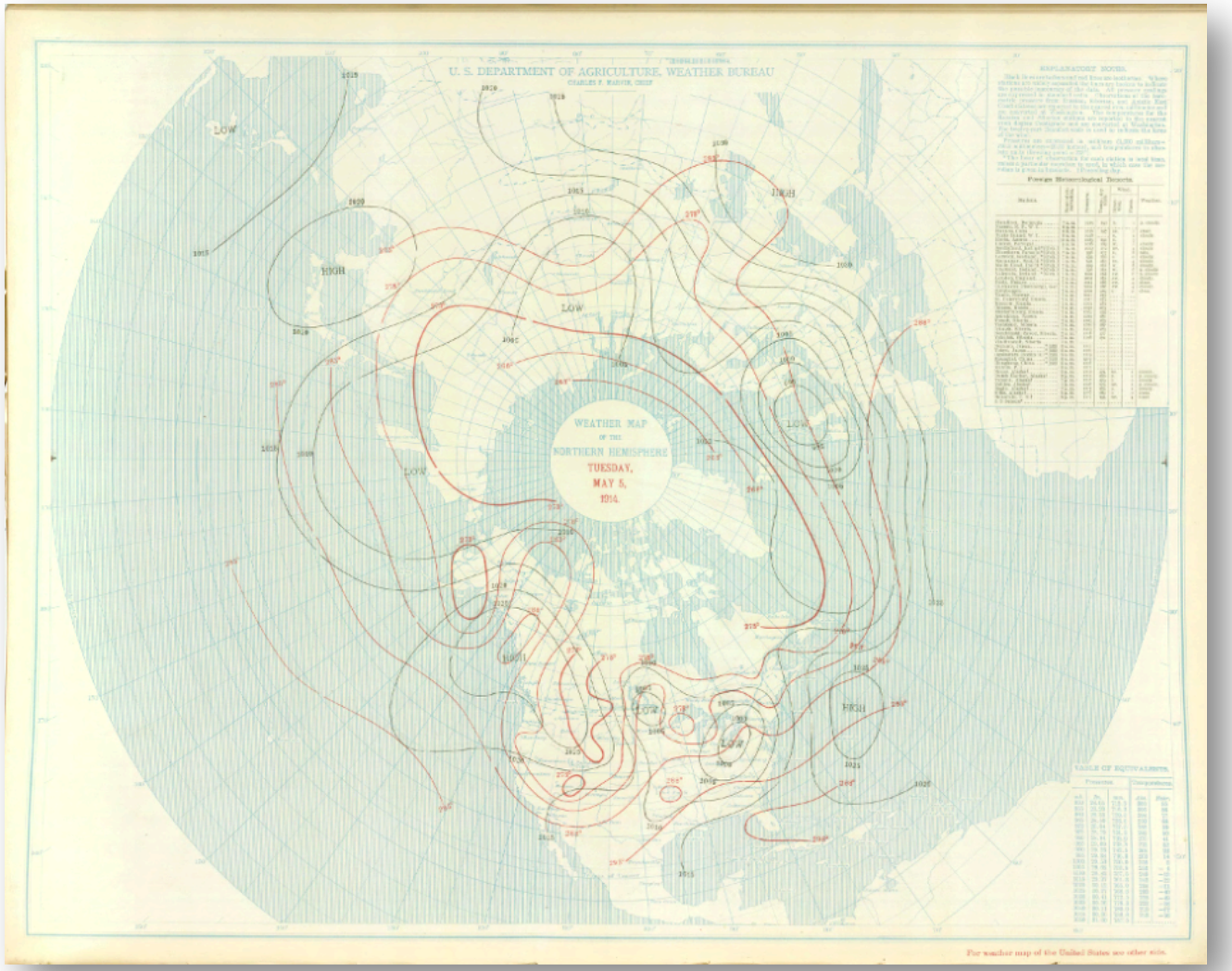
Climate models and climate data

Data friction

Climate controversies

Conclusion

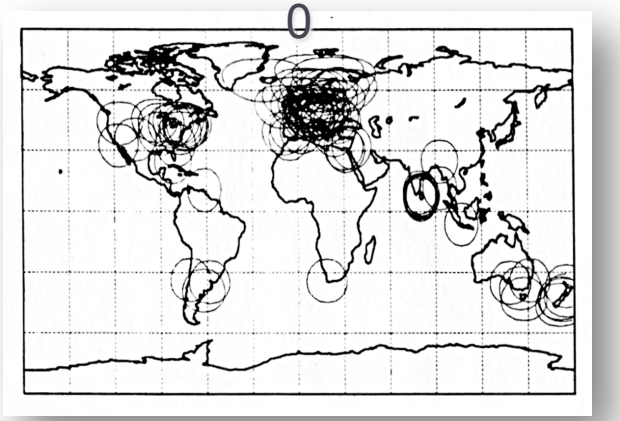
1872 War Dept. weather map



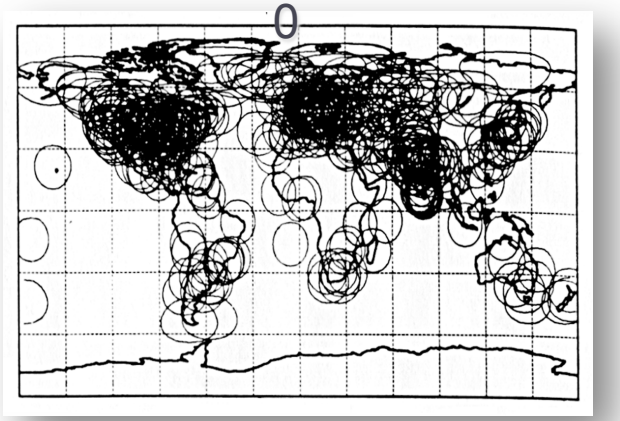
Synoptic map of the northern hemisphere, 5 May 1914

Global coverage by surface stations

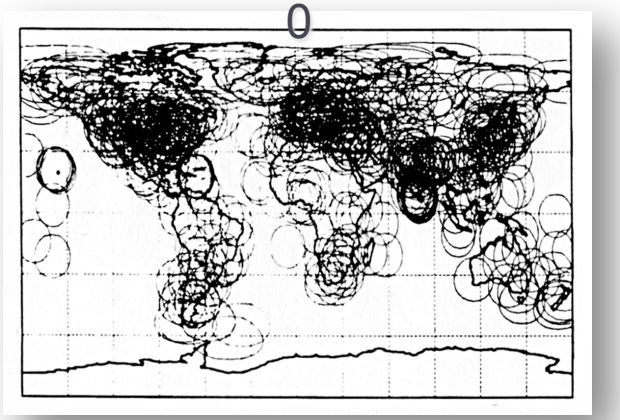
1870



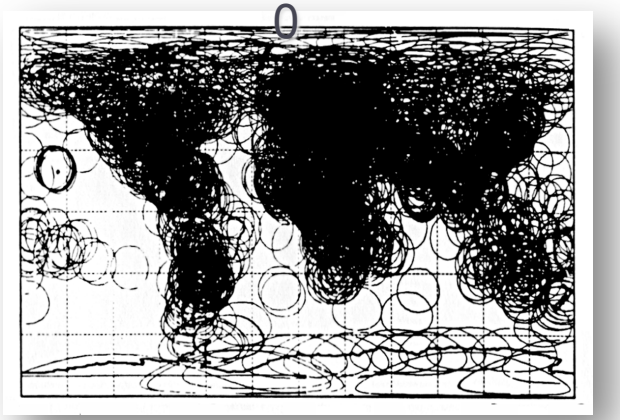
1900



1930



1960



Source: Hansen & Lebedeff (1987), "Global Trends of Measured Surface Air Temperature," *JGR*

L.F. Richardson's "forecast-factory" (1922)

Illustration by François Schuiten (1990s)

Introduction

Weather forecasting

Computers change everything

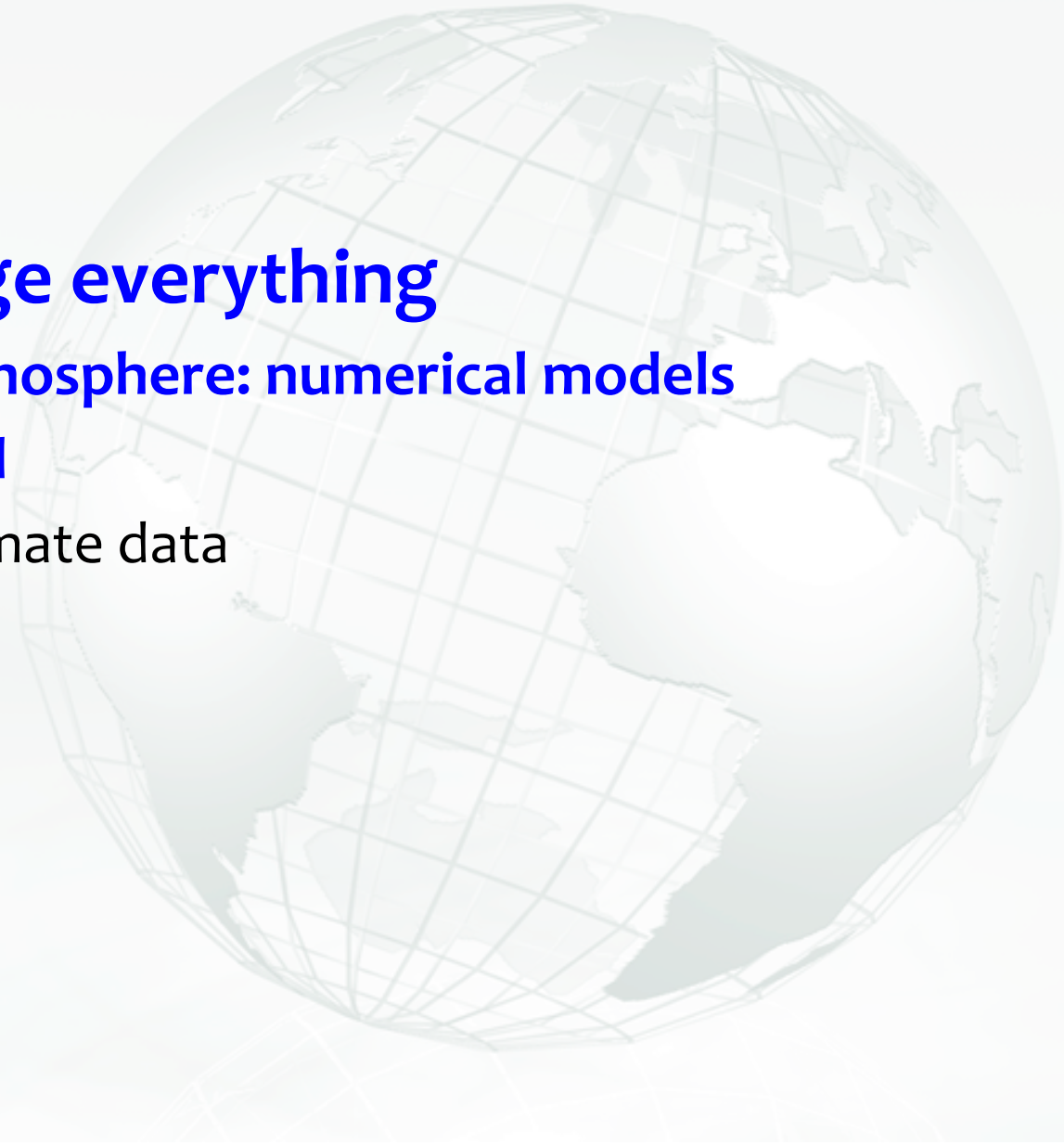
- **simulating the atmosphere: numerical models**
- **making data global**

Climate models and climate data

Data friction

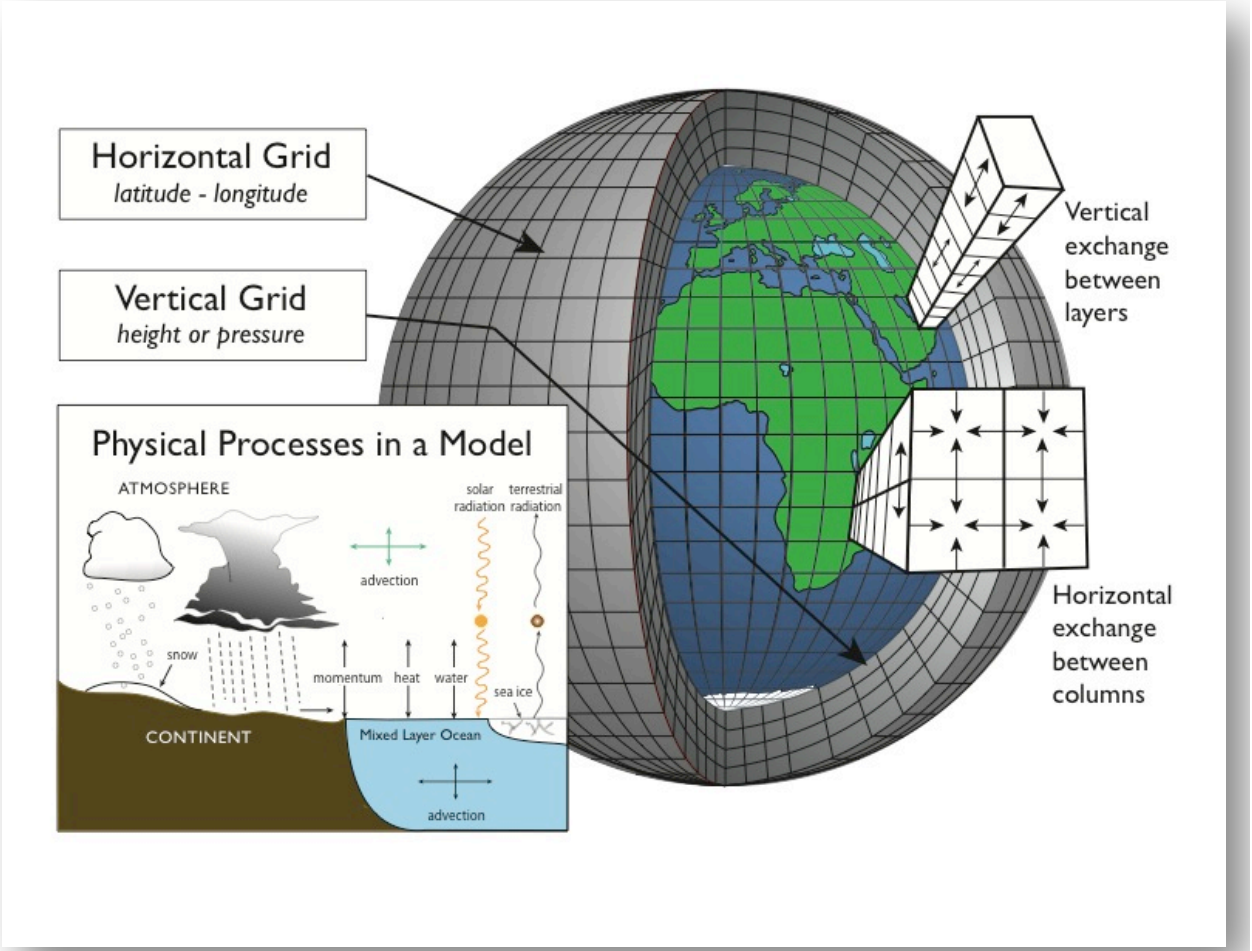
Climate controversies

Conclusion

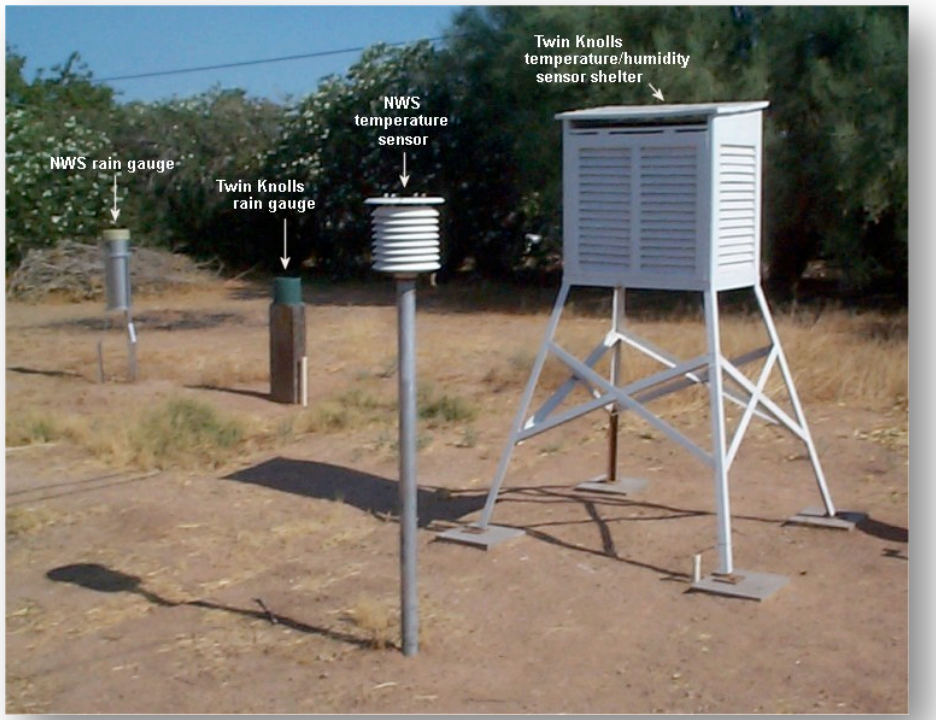


Numerical weather prediction models

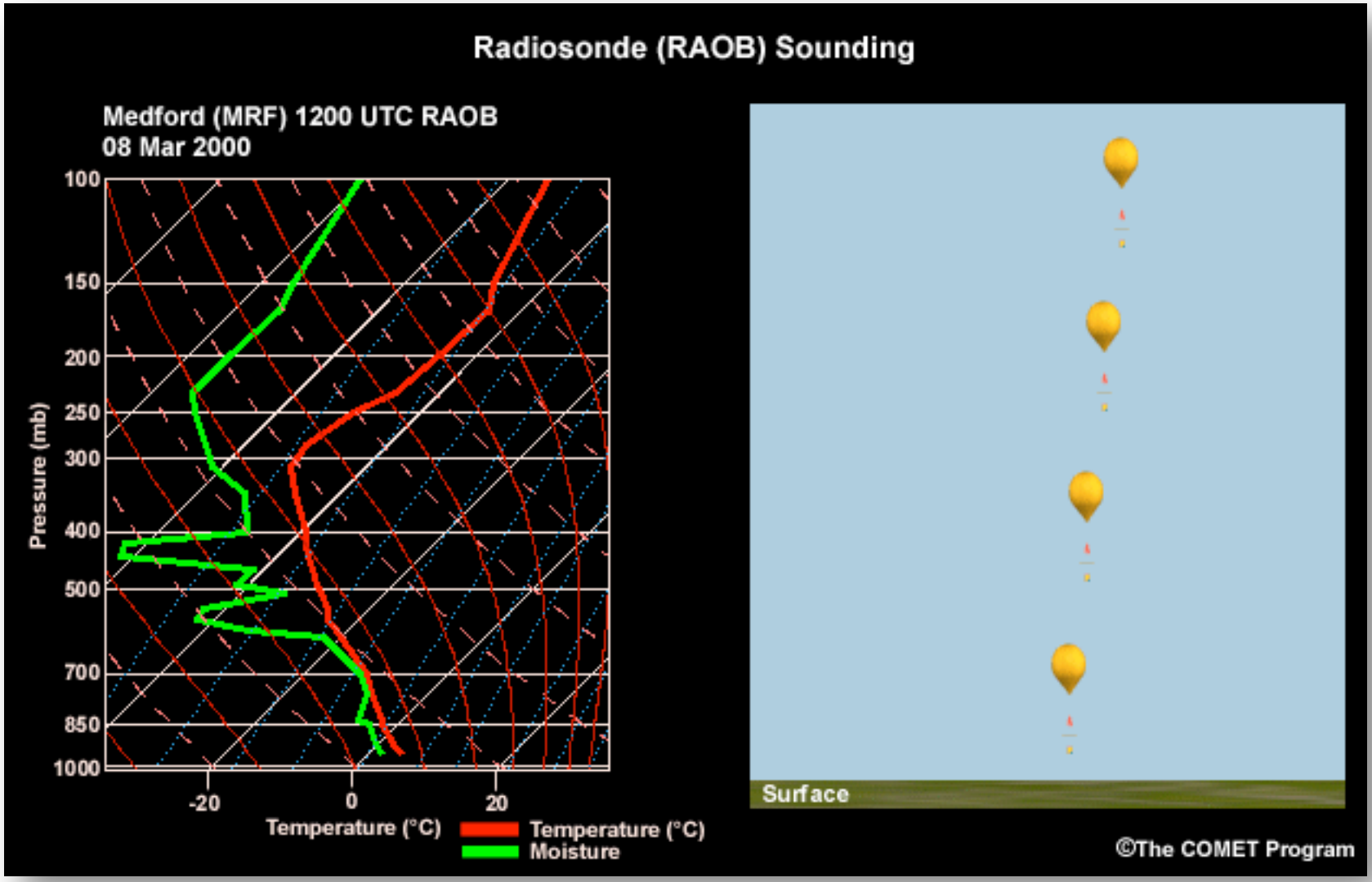
- Operational since 1955
- Simulate evolution of initial state (observations)
- Forecast models drove need for more data
  - How to fill empty gridpoints?
  - How to reconcile heterogeneous data sources?



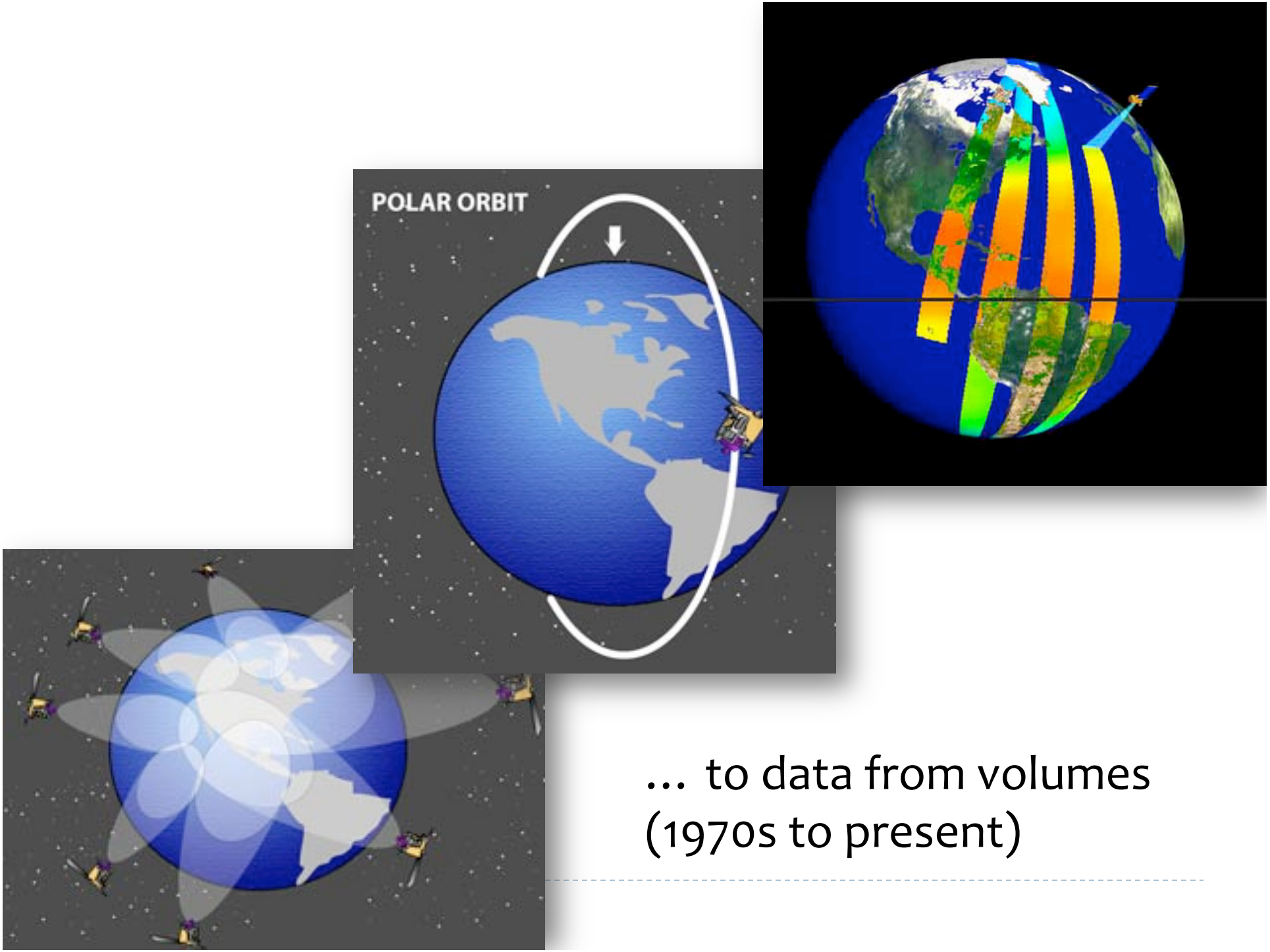
Data from points
(1600s to present)…


Twin Knolls
temperature/humidity
sensor shelter
NWS
temperature
sensor
NWS rain gauge
Twin Knolls
rain gauge

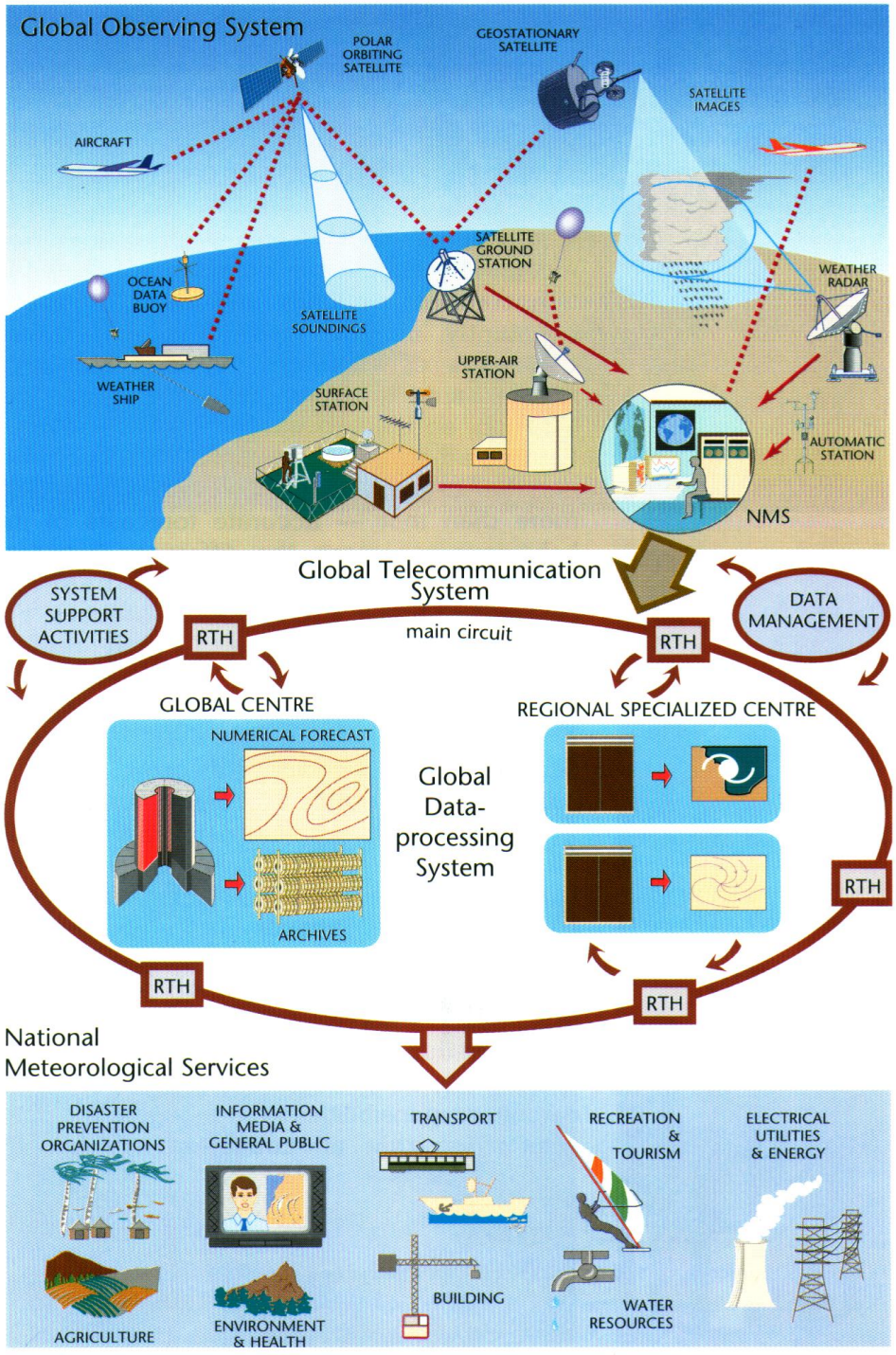

NOAA



… and data from lines (1950s to present)…



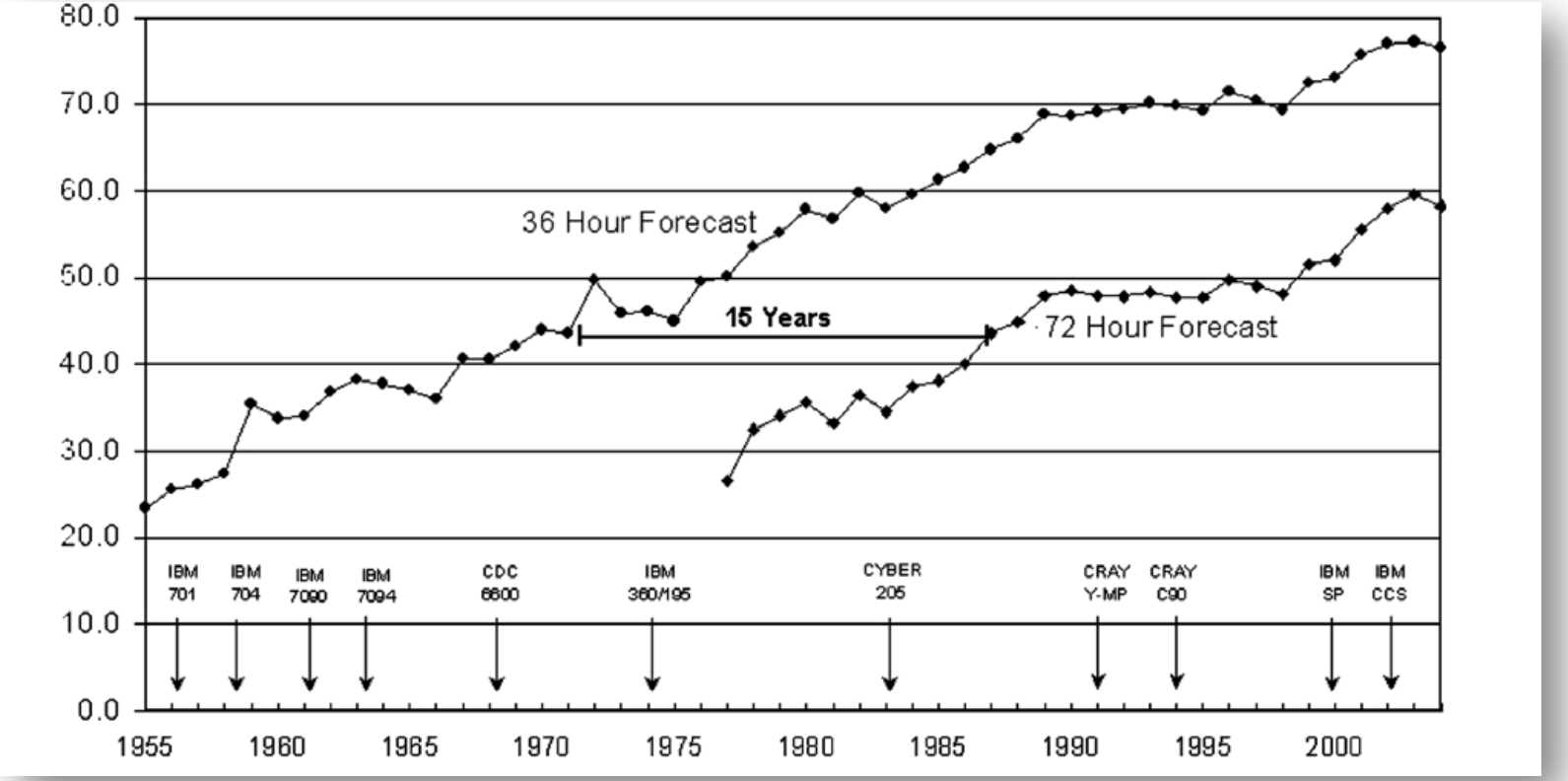
… to data from volumes
(1970s to present)



World Weather Watch

- initial planning early 1960s
- operational about 1968

Forecast skill improvement: US Weather Service, 1955-2010



Forecast quality, expressed as a percentage of a "perfect" forecast

Source: Peter Lynch, "The Origins of Computer Weather Prediction and Climate Modeling," *Journal of Computational Physics* 227, no. 7 (2008): 3431-44

Introduction

Weather forecasting

Computers change everything

Climate data and climate models

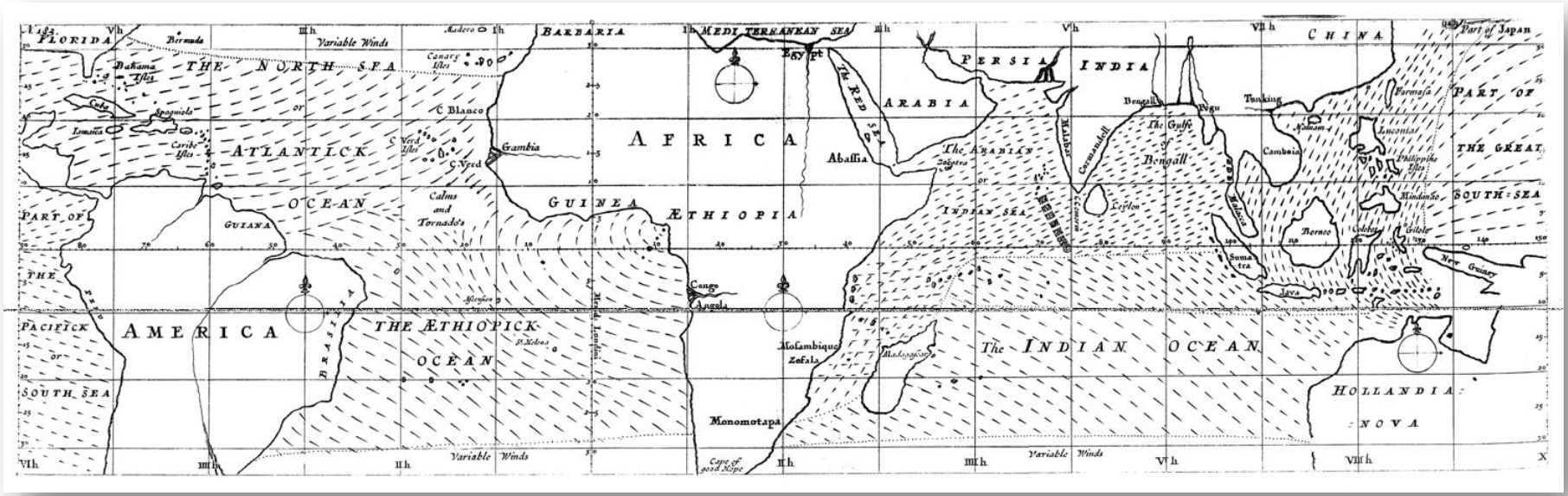
- **understanding global circulation**
- **making global data**
- **climate simulations**

Data friction

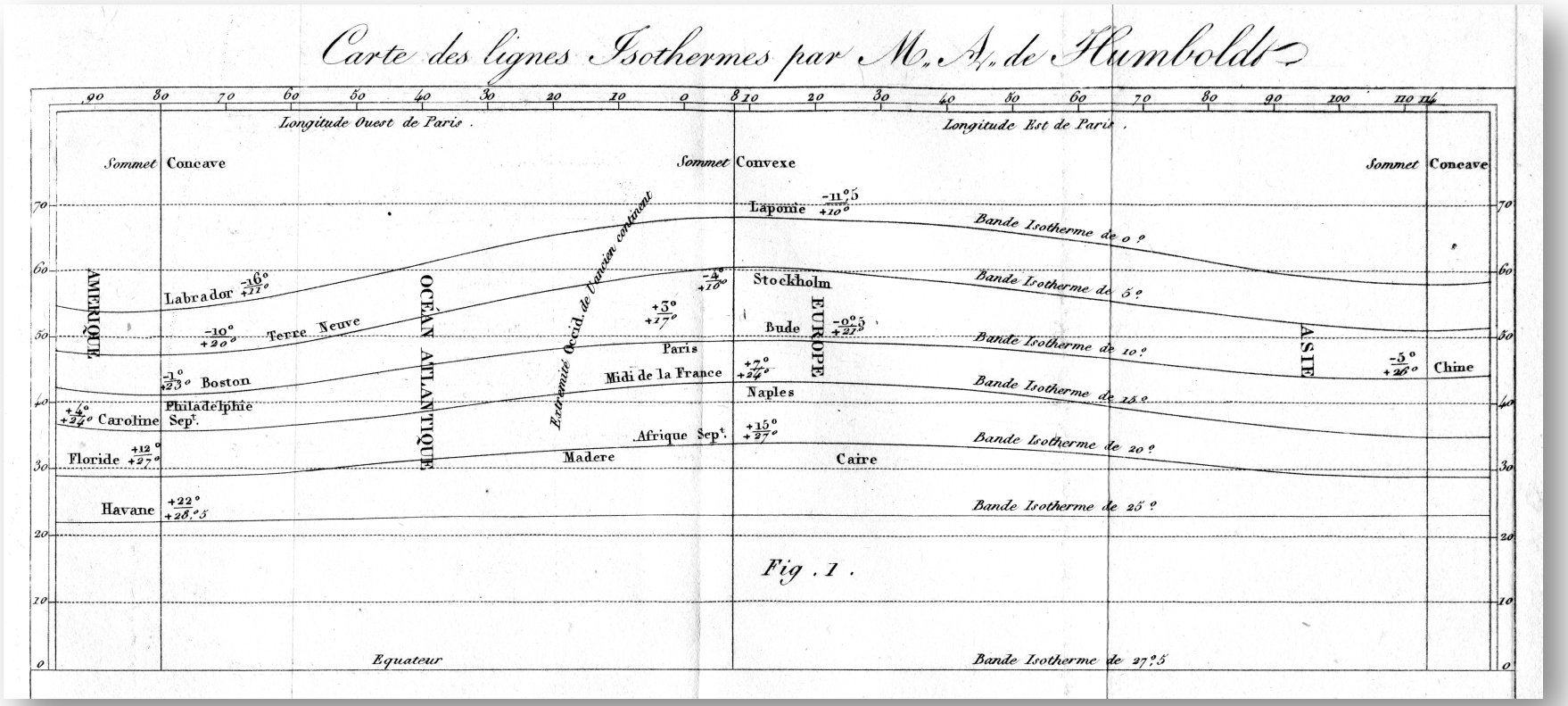
Climate controversies

Conclusion

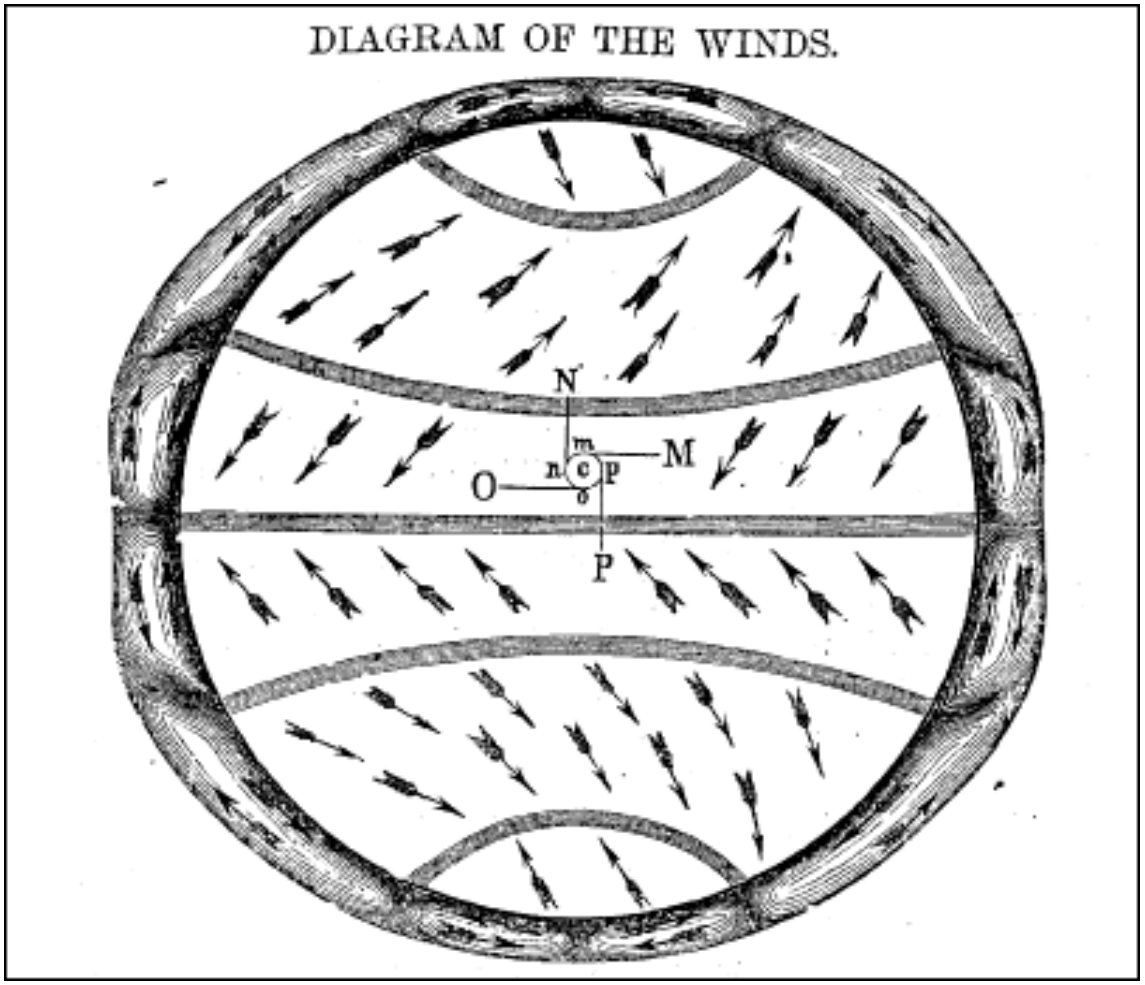
The trade winds (Halley 1686)



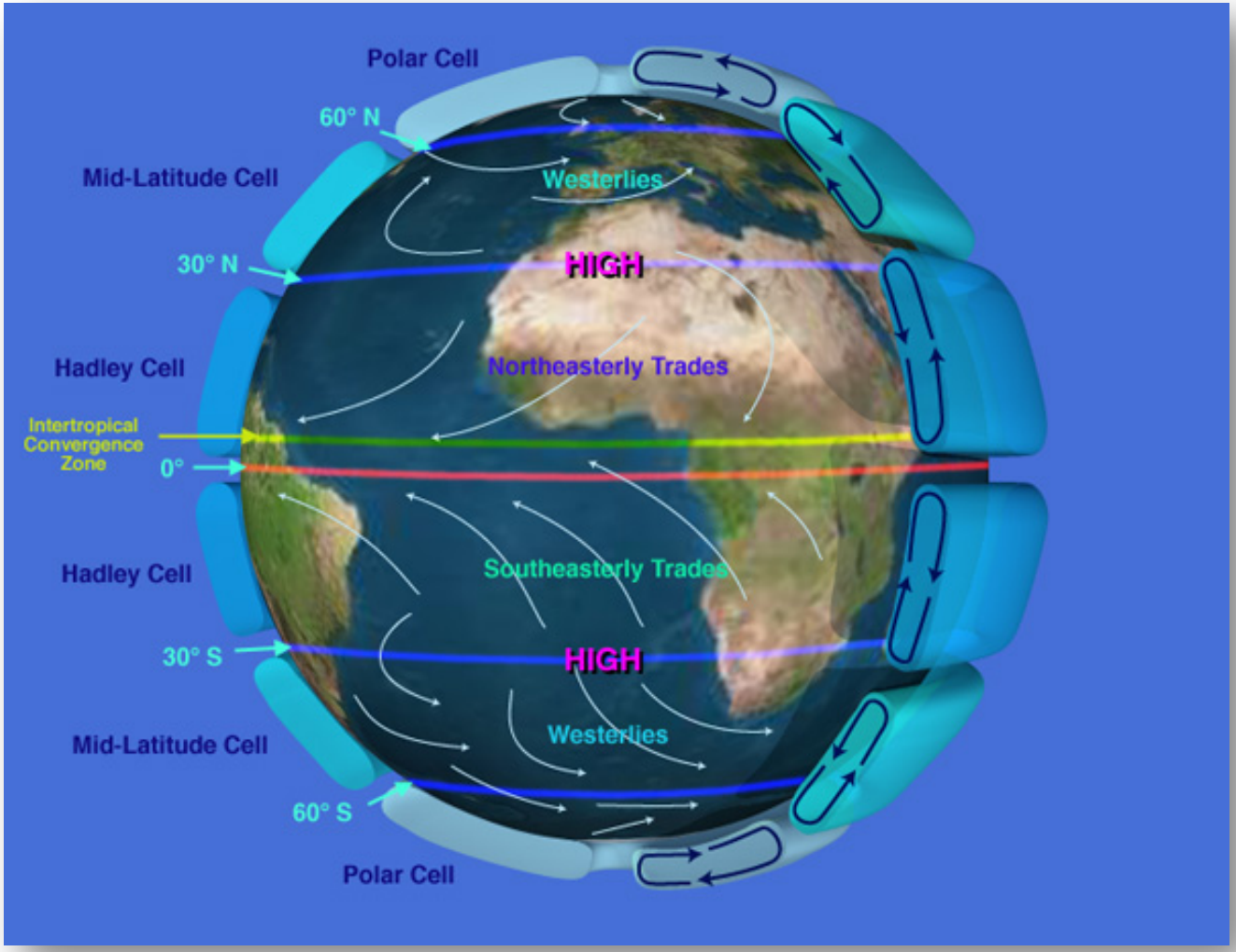
Isotherms (von humboldt, 1817)



The global circulation (Ferrel 1860)

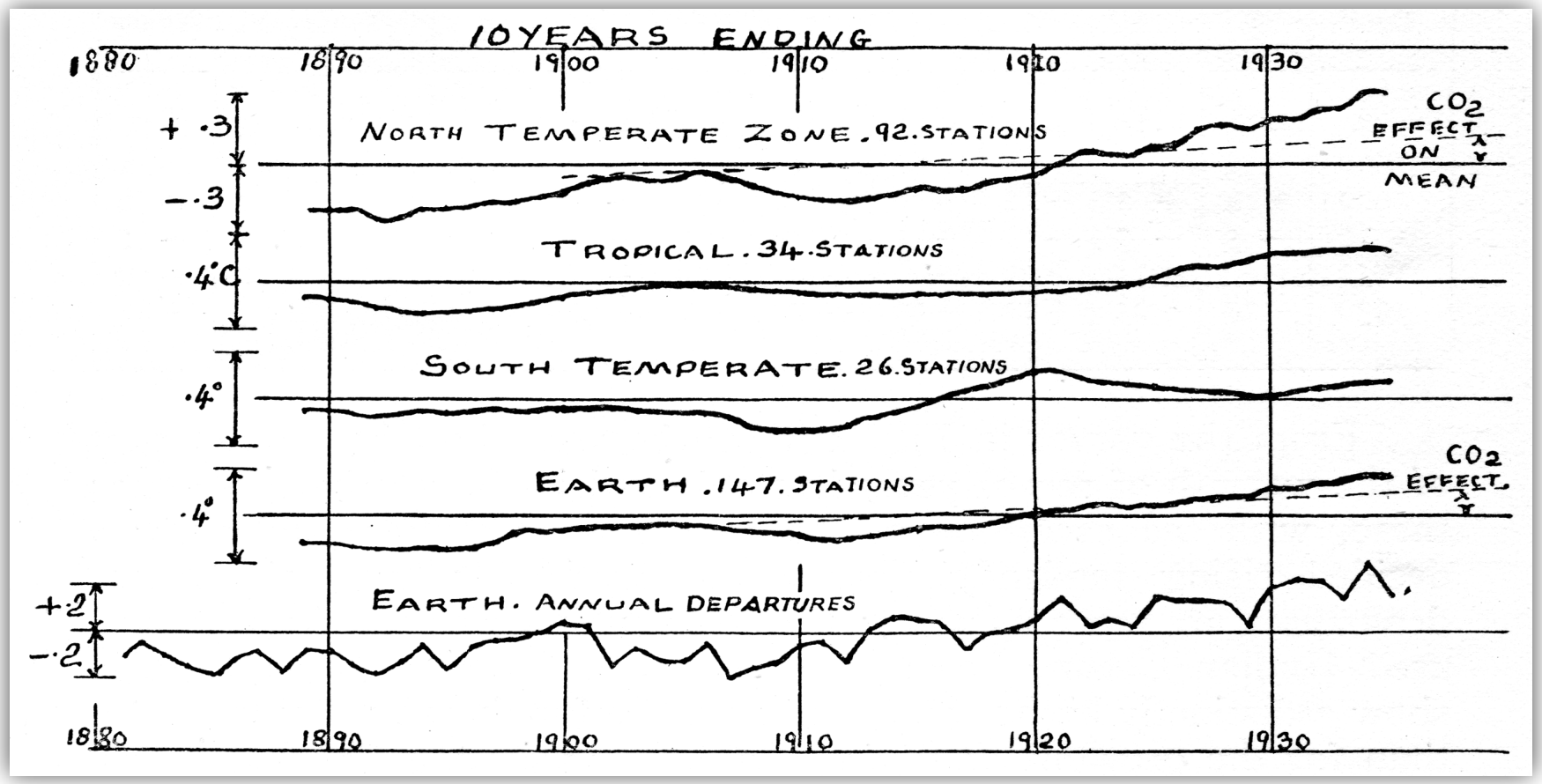


The global circulation: a recent view



Callendar 1938: is CO_2 causing global warming?

Global temperature from ~150 stations, 1890-1937



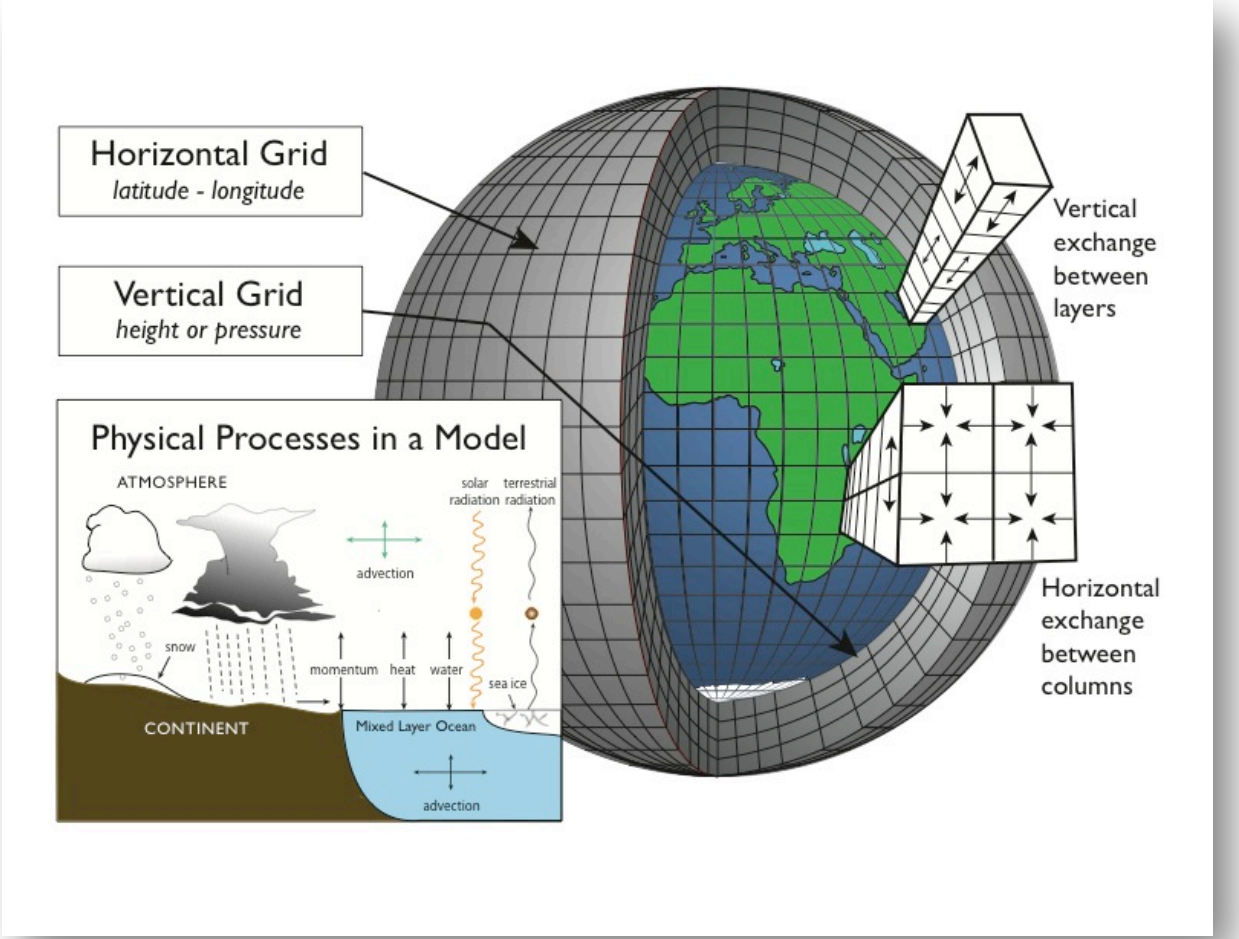
Callendar 1938: is CO_2 causing global warming?

Discussion at the Royal Meteorological Society

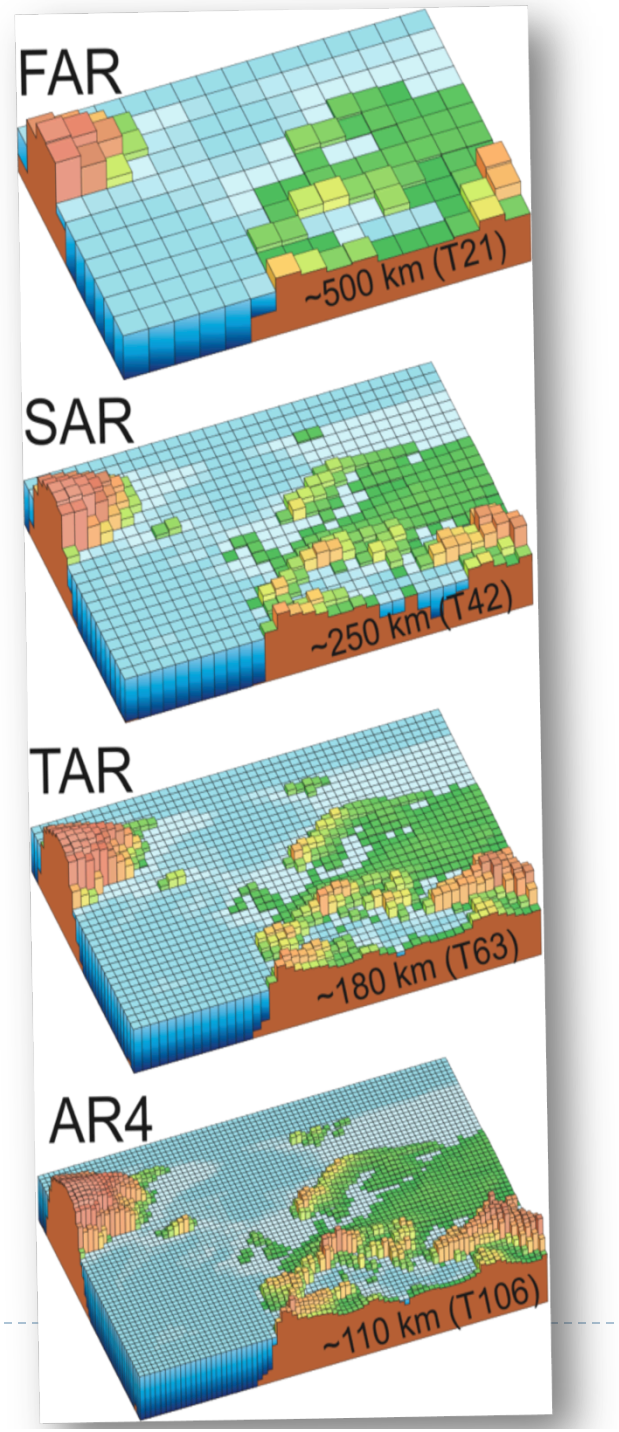
- "Coincidence…. one of the peculiar variations in all meteorological elements" — G. Simpson
- Temperature increase explained by changes in general circulation — C.E.P. Brooks
- Change in heat absorption would produce complex circulation effects, not simple temperature rise — D. Brunt

Computer models of the atmosphere

- Simulations of the climate system
  - Run to equilibrium (initial state unimportant)
  - Statistical properties, not day-to-day weather, are of interest
- Problem: where to get data for validating climate models?



Increasing resolution in climate simulations



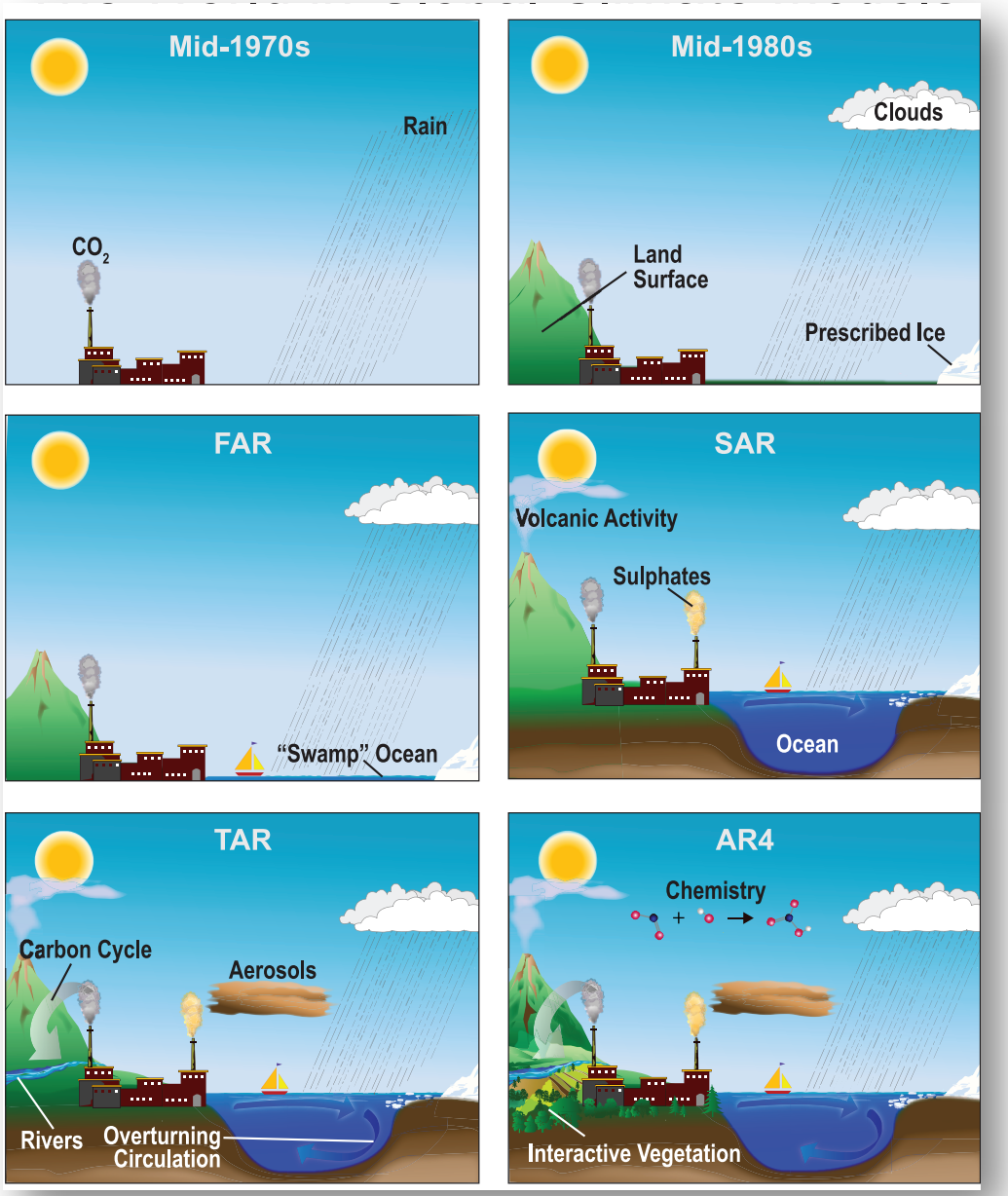
1990

1996

2000

2007

Physical processes included in climate models



Climate sensitivity from GCMs

Assessment	Range of GCM results (°C)	Equilibrium climate sensitivity	"Best guess" (°C)
NAS 1979	2-3.5	1.5-4.5	3
NAS 1983	2-3.5	1.5-4.5	3
Villach 1985	1.5-5.5	1.5-4.5	3
IPCC 1990	1.9-5.2	1.5-4.5	2.5
IPCC 1992	1.7-5.4	1.5-4.5	2.5
IPCC 1994	not given	1.5-4.5	2.5
Bolin 1995	not given	1.5-4.5	2.5
IPCC 1995	1.9-5.2	1.5-4.5	2.5
IPCC 2001	2.0-5.1	1.5-4.5	2.5
IPCC 2007	2.1-4.4	2-4.5	3
IPCC 2013	2.1-4.7	1.5-4.5	not given

Introduction

Weather forecasting

Computers change everything

Climate data and climate models

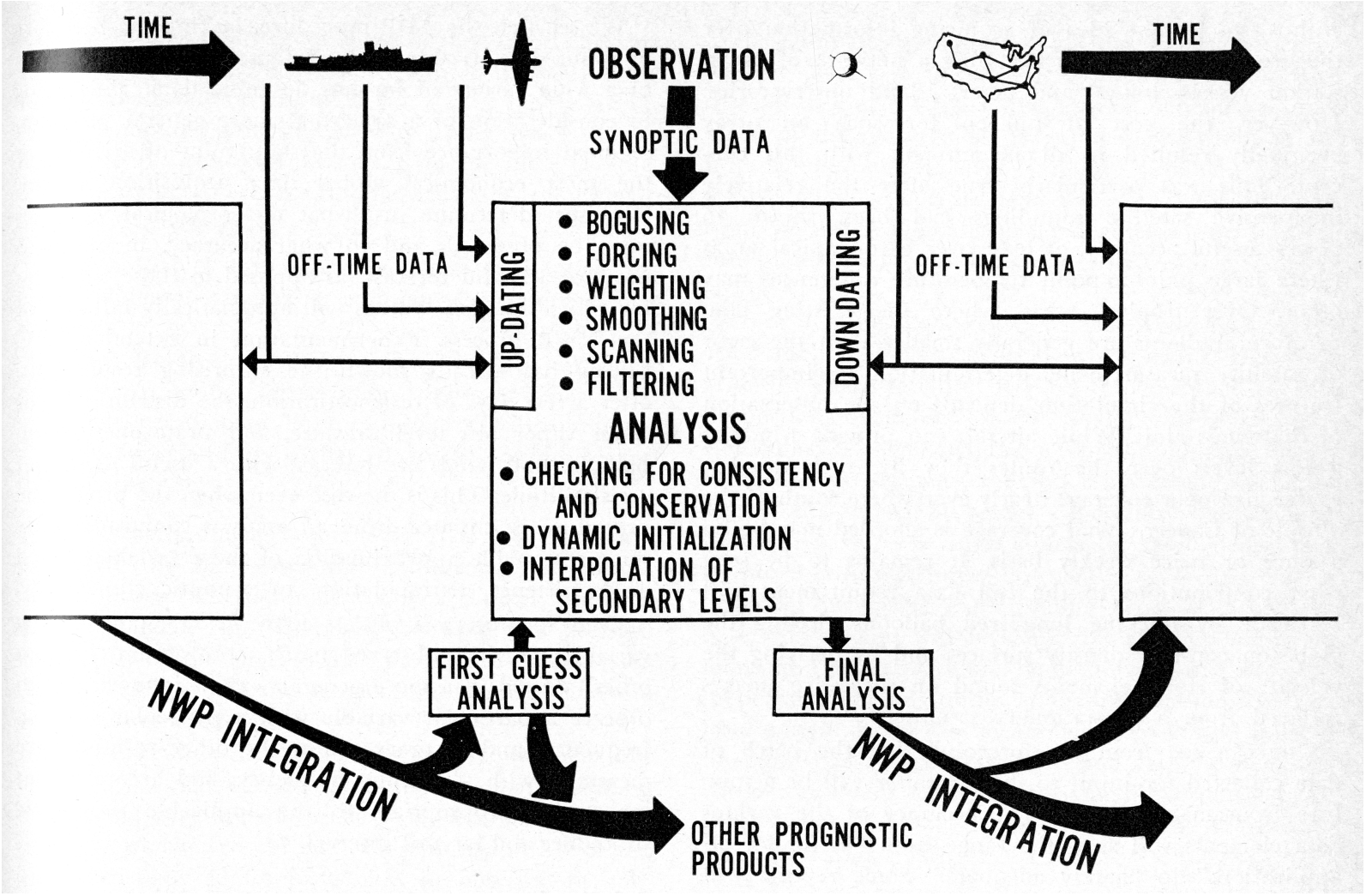
Data friction

- **making data global: versions of the past**
- **infrastructural inversion as a method**

Climate controversies

Conclusion

Data assimilation: 1960s

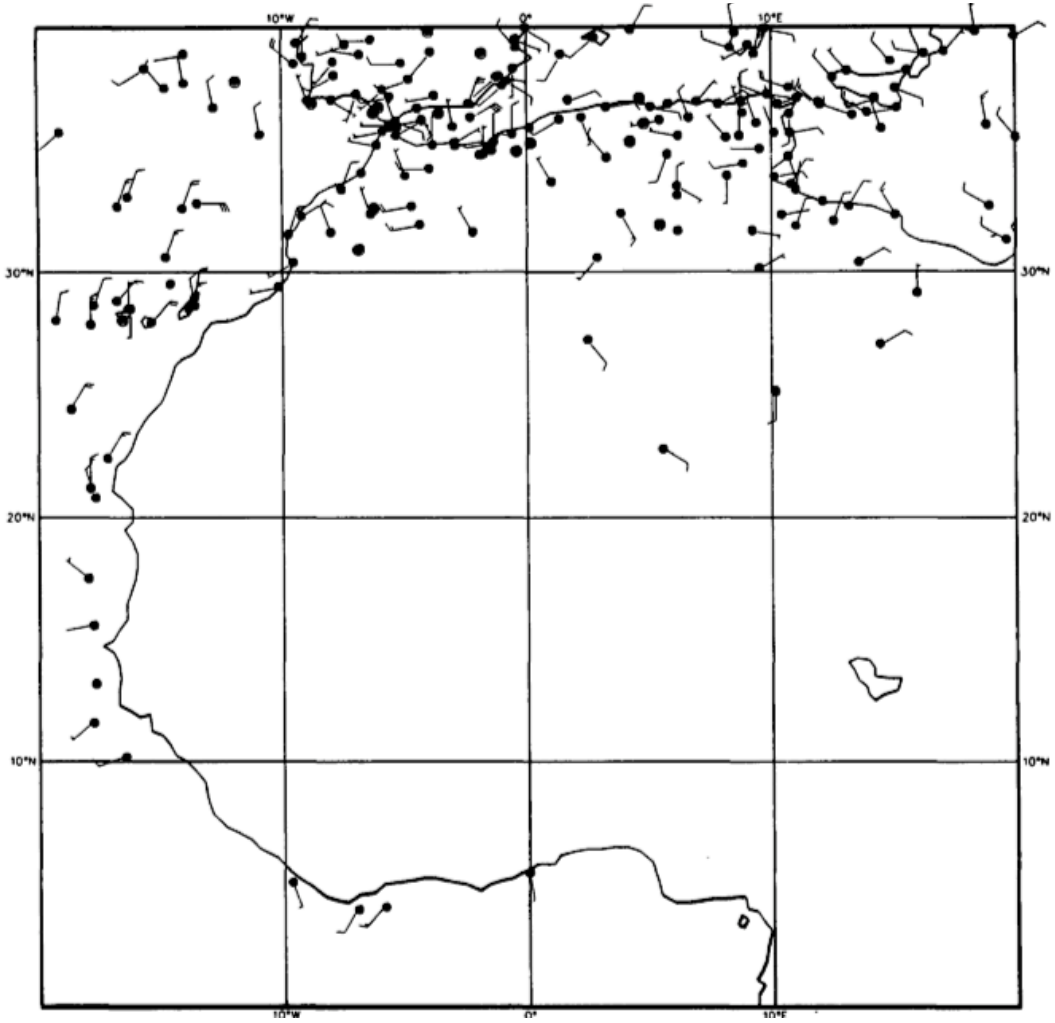


Simulated data are better than real data

- Observations are samples of huge volume
- Model-produced data have much higher resolution than observations

"A realistic global [analysis] model can be viewed as a unique and independent observing system that can generate information at a scale finer than that of the conventional observing system" (Bengtsson & Shukla 1988)

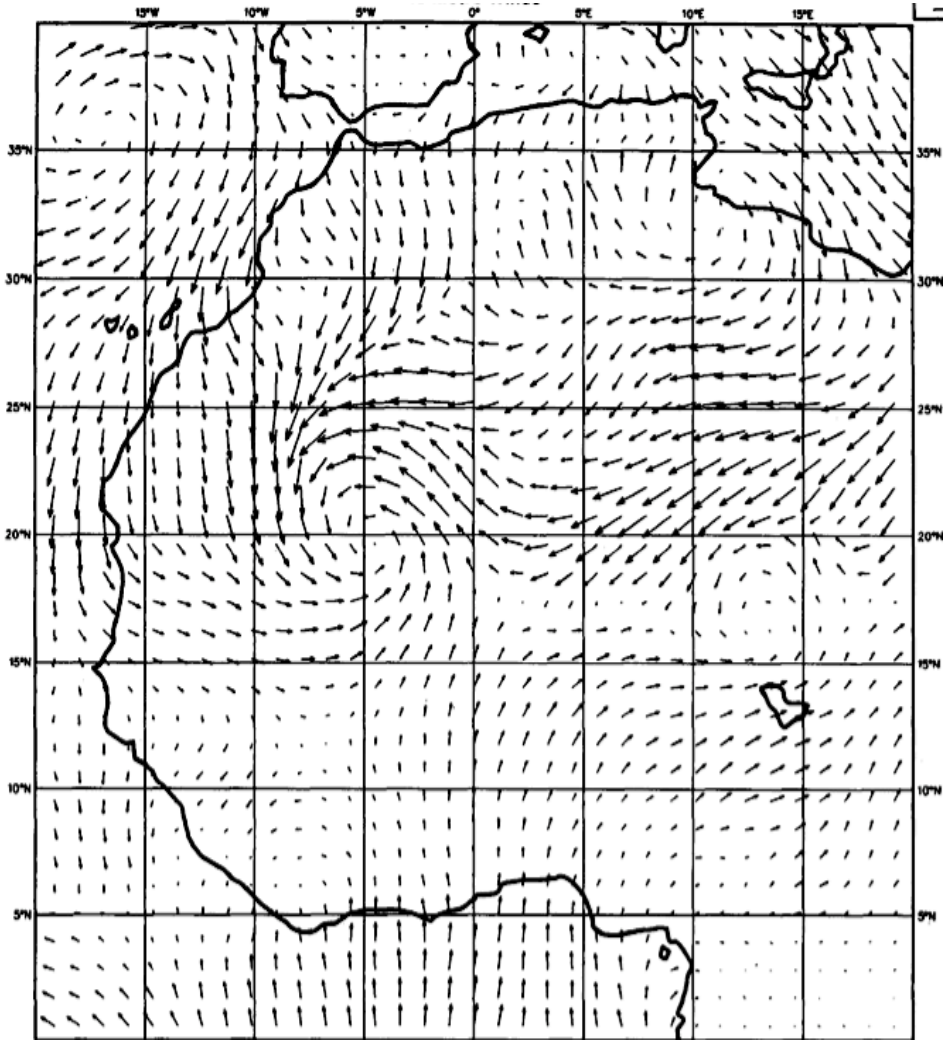
Surface Wind Obs, 28 Aug 1985



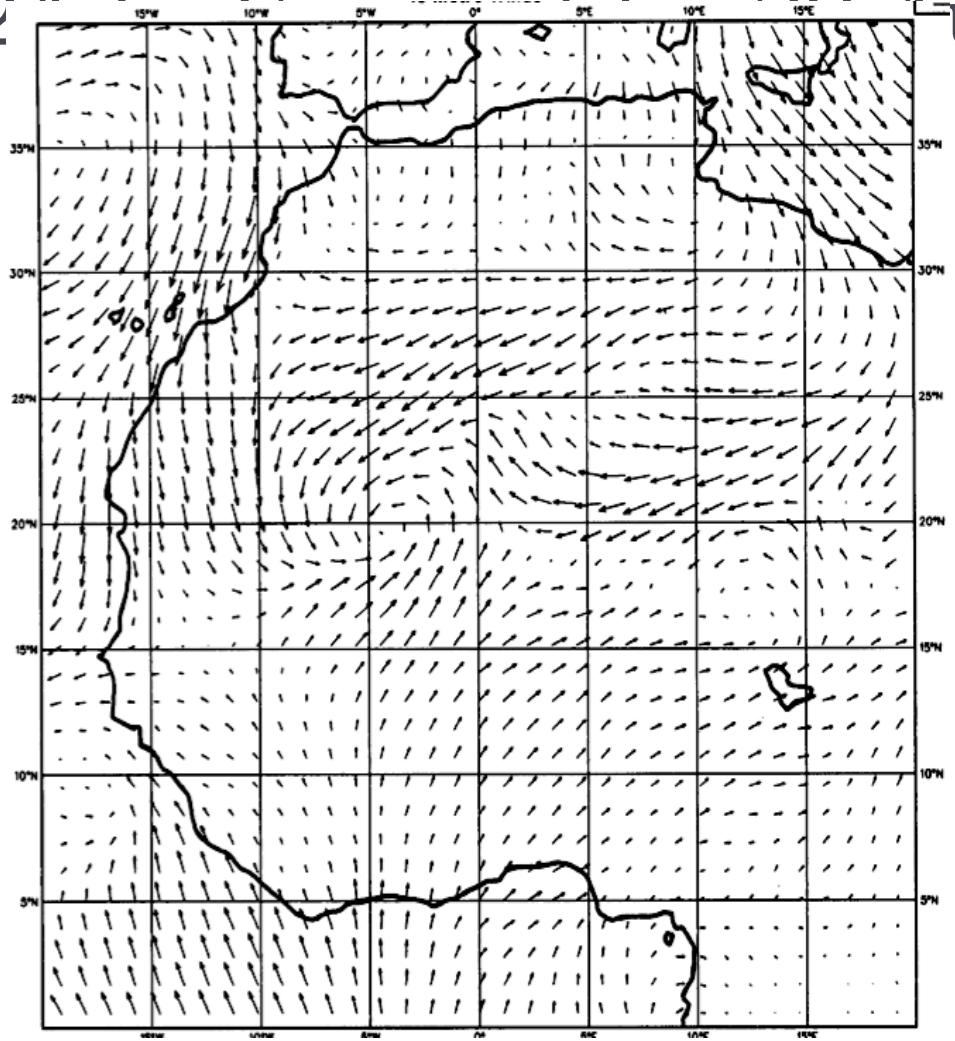
25-30 Aug 1985: breakdown in Global Telecomm. System

Almost no N. African data were transmitted

ECMWF Analysis, 28 Aug 1985



ECMWF 24-h Forecast for 00 Aug 1985



Meteosat Photo — N. Africa, 29 Aug 1985

Data friction

- Data are *things*
- Collect, process, move, store, manage, provide access…

Entry hall, US National Weather Records Center, Asheville NC, early 1960s

Country	*Methods used to calculate mean daily temperature*
Egypt	1/2(max + min); means of 3-hourly observations, 1/8(00 +03 + …21); 1/4(09 + 21 + max + min); 1/4(06*Z* + 12*Z* + 18*Z* + min); means of 24 hourly values (exact hours unknown)
France	1/2(max + min); 1/3(06 + 13 + 21); 1/3(06 + 14 + 22); means of 24 hours, 1/24(01*Z* + 02*Z* +... 24*Z*); means of 8 3-hourly observations
Ghana	1/8(03 + 06 + …24); 1/2(max + min)
Guyana	1/2(max + min); 1/12(00 + 02 + ...22); 1/3(07 + 13 + 18) local time; 1/2(12Z + 18Z).
Tunisia	Means of 24 hours (exact hours unknown); 1/2(max + min); 1/4(07 + 13 + 19 + (19 + min)/2)
U.S.S.R.	1/4(01 + 07 + 13 + 19); 1/3(07 + 13 + 21); 1/4(01 + 07 + 13 + 21); 1/4(07 + 14 + 21 + 21) 105°E meridian time; means of 2-4 daily observations in 53 different combinations

Source: Palutikof and Goddess, 1986

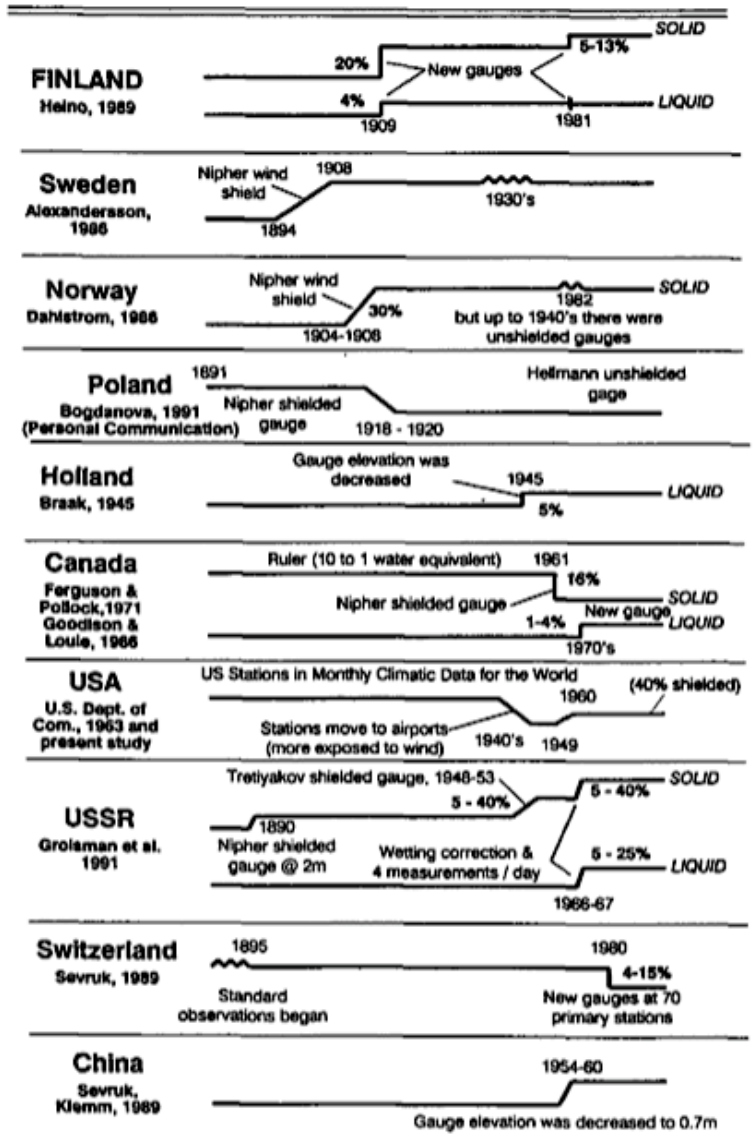


FIG. 1. Some important precipitation discontinuities over the last 100 years at many primary observing stations within various mid- and high-latitude Northern Hemisphere countries.

Changes in instrumentation (Karl et al. 1993)

Changes in standard observing hours

(Karl et al. 1993)

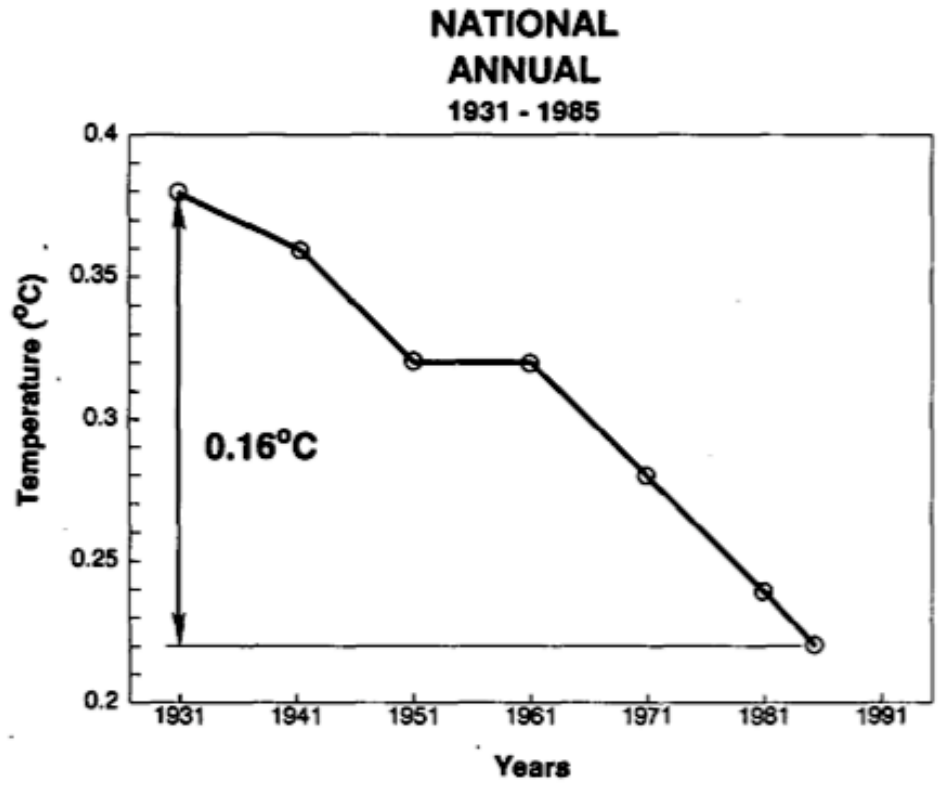
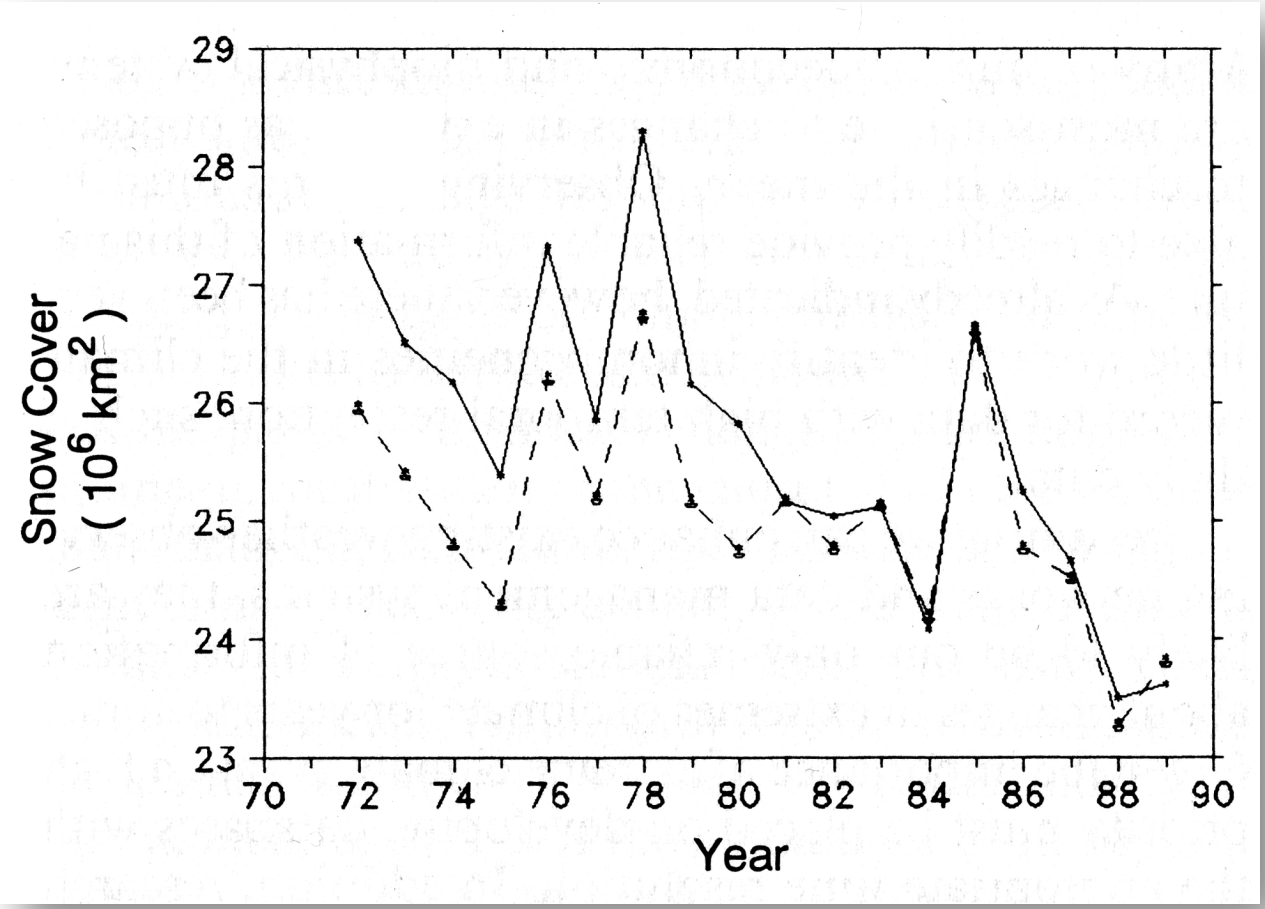


FIG. 3. Estimated effects of changes in observation time on the contiguous United States area-average temperature based on all stations that measure temperature relative to a calendar-day observing schedule.

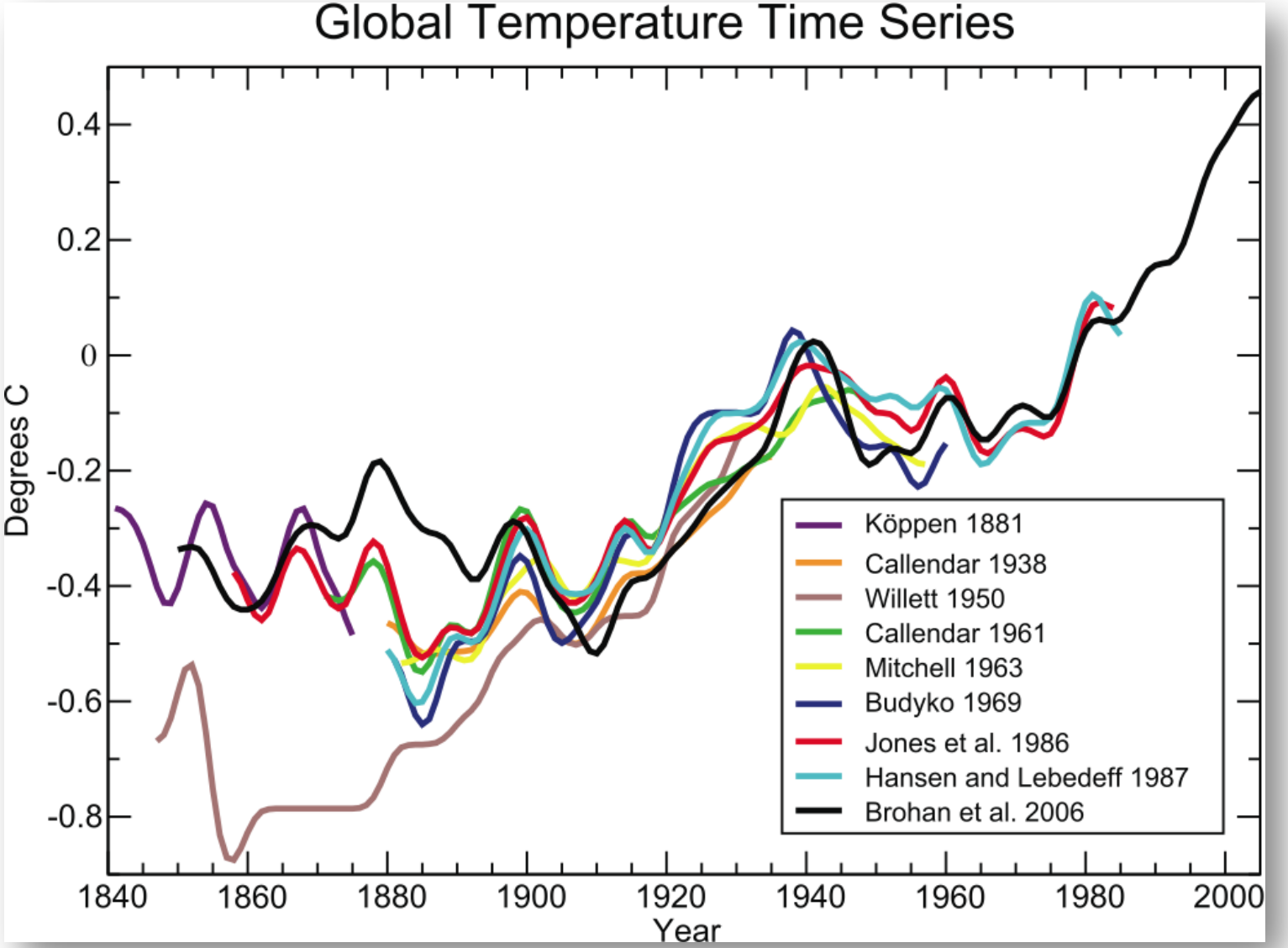
Changes in data models



Northern hemisphere snow cover from NOAA polar orbiting satellites, processed using a consistent data reduction algorithm (solid line) vs. the same data as processed by earlier algorithms (dashed line).

Infrastructural inversion as a method

- Exposing the infrastructure to find and fix data problems
- … a fundamental method in climate science

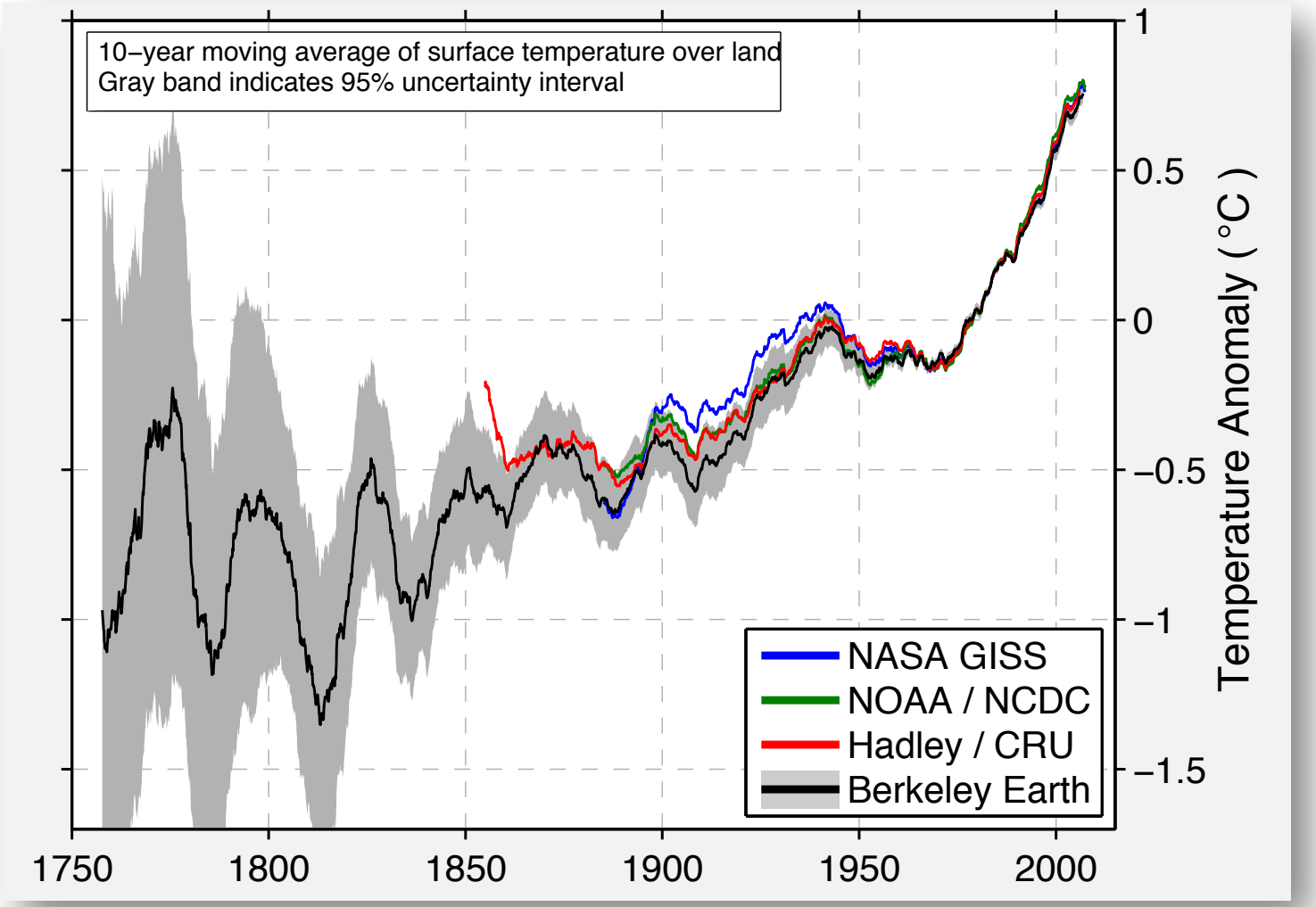


Source: IPCC AR4, 2007

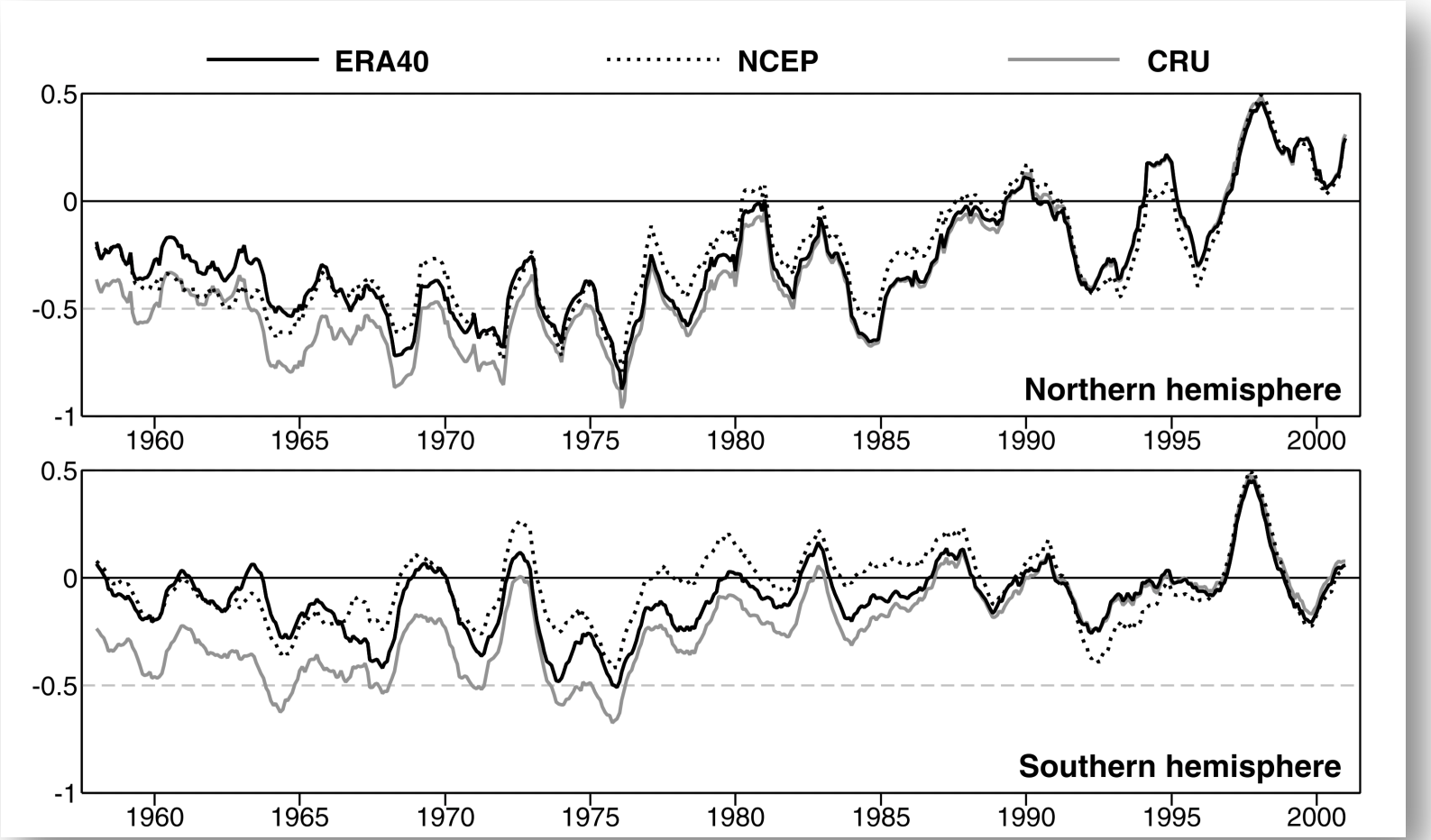
Infrastructural inversion at work:
Making data global

- Köppen 1881: fewer than **100 stations**
- Callendar 1938: about **200 stations**
- Willett 1950: **183 stations**
- Callendar 1961: **450 stations**
- Mitchell 1963: **183 stations**
- Jones et al. 1986: **2194 stations**
- Brohan et al. 2006: **4349 stations**
- Muller et al. 2012 : **39,340 stations**

Berkeley Earth (2012 published, after peer review)
Global land surface average temperature



Making global data *and* making data global:
Reanalysis (ERA, NCEP) vs. observations



Introduction

Weather forecasting

Computers change everything

Climate data and climate models

Data friction

Climate controversies

- satellite data
- surfacestations.org

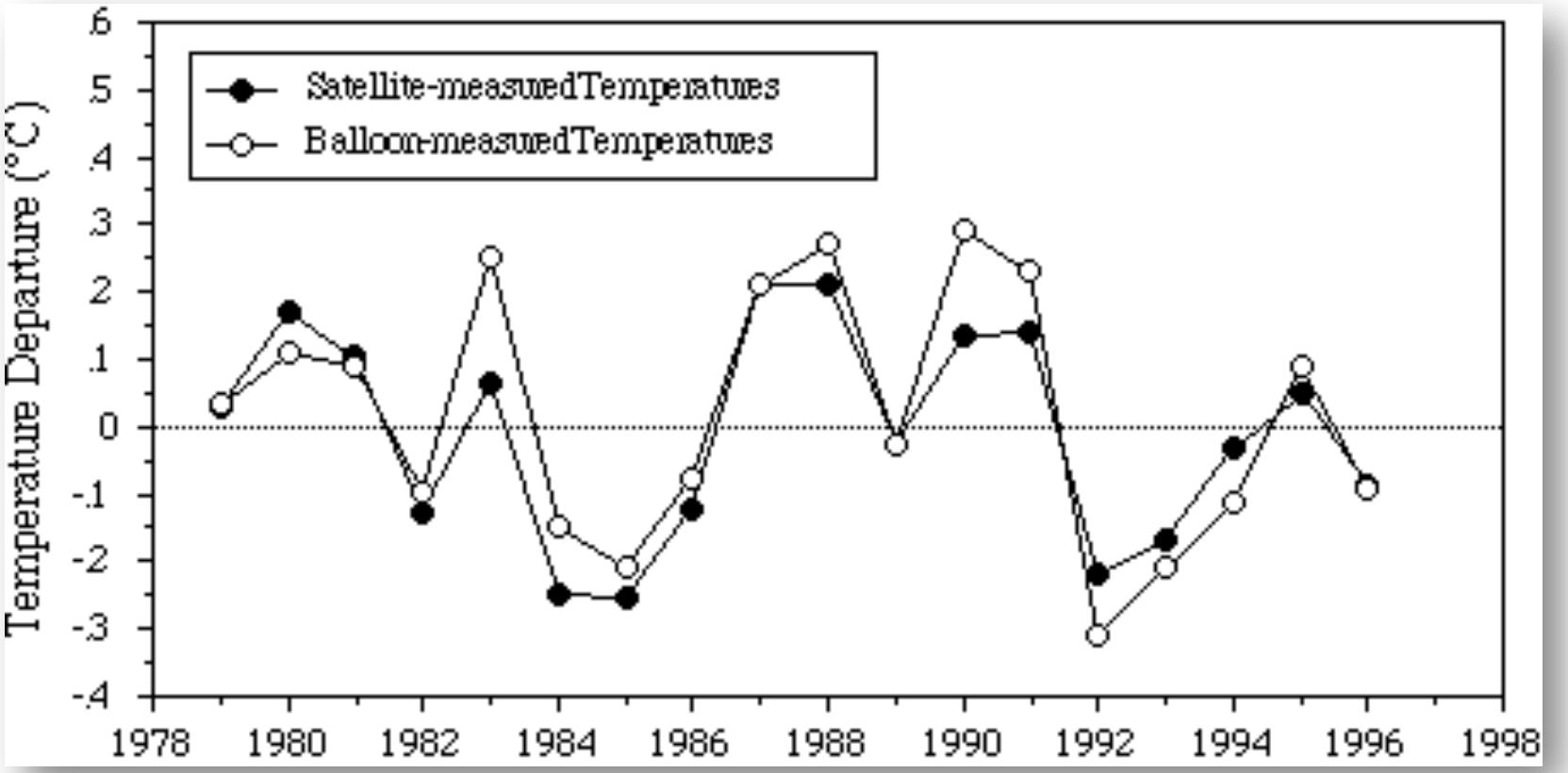
Conclusion

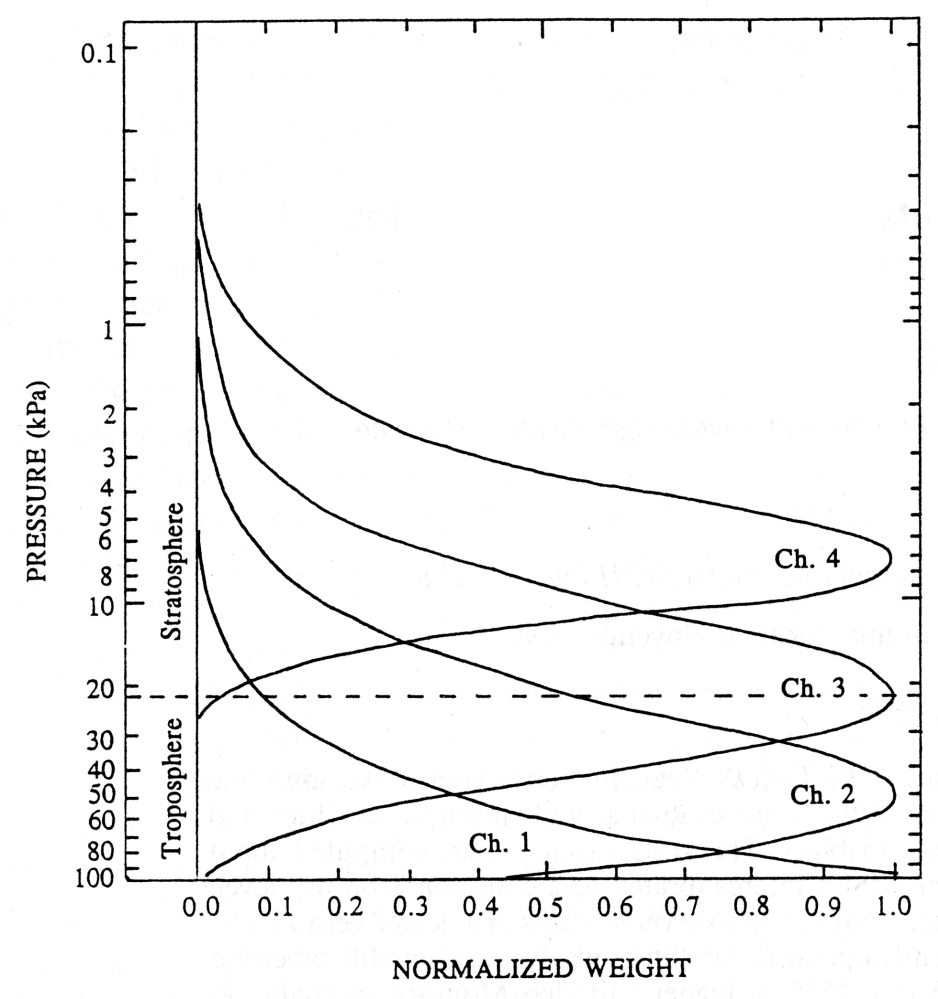
The Scientific Integrity Hearings, 1995

"Hurrell and Trenberth estimate the temperature of the atmosphere through a simple linear regression model…, and a global climate model simulation …, [but] *the MSU data actually measure the temperature of the free atmosphere.*"

— John Christy, testimony to House Science Committee

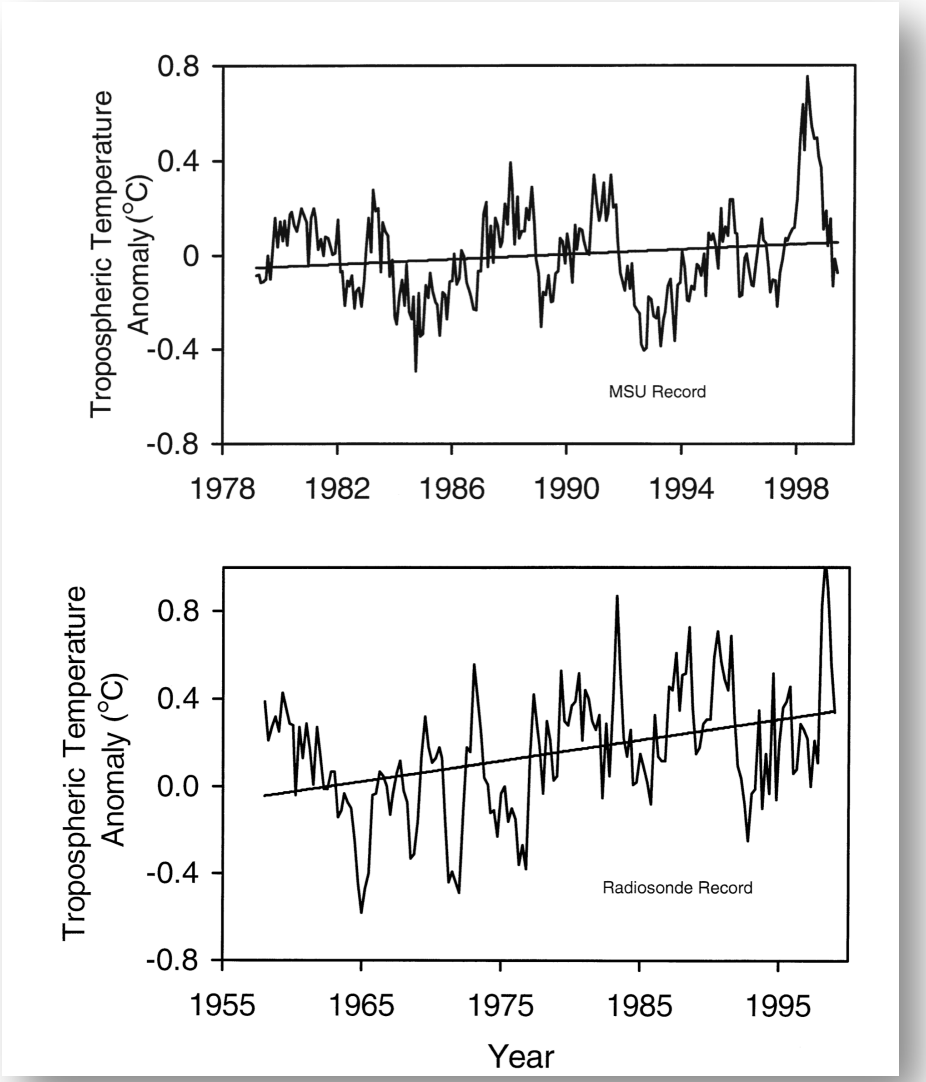
Satellites vs. radiosondes (1996)





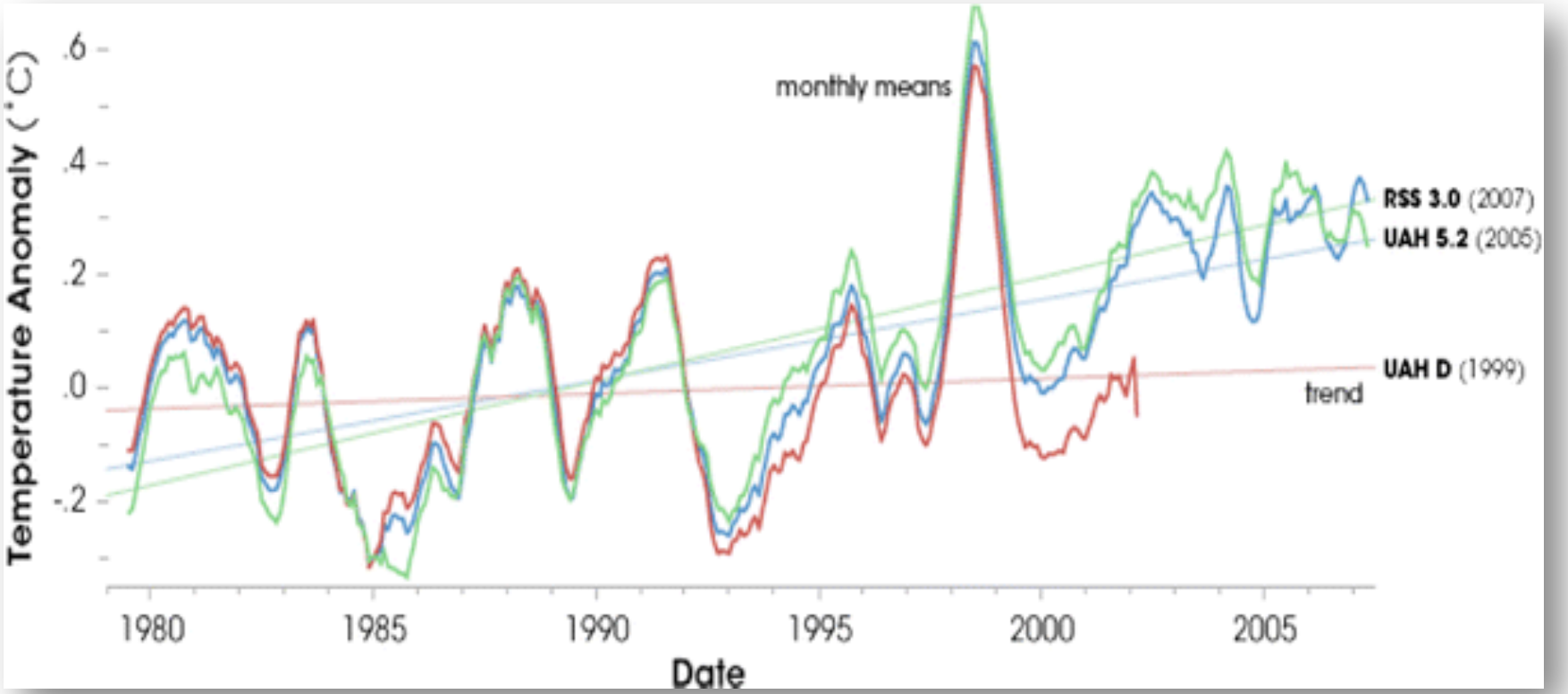
Weighting functions for Microwave Sounding Unit channels 1-4, used to create a vertical atmospheric temperature profile from top-of-atmosphere radiances in four spectral bands (channels)

Tilting the line: satellites vs. radiosondes ca. 1999



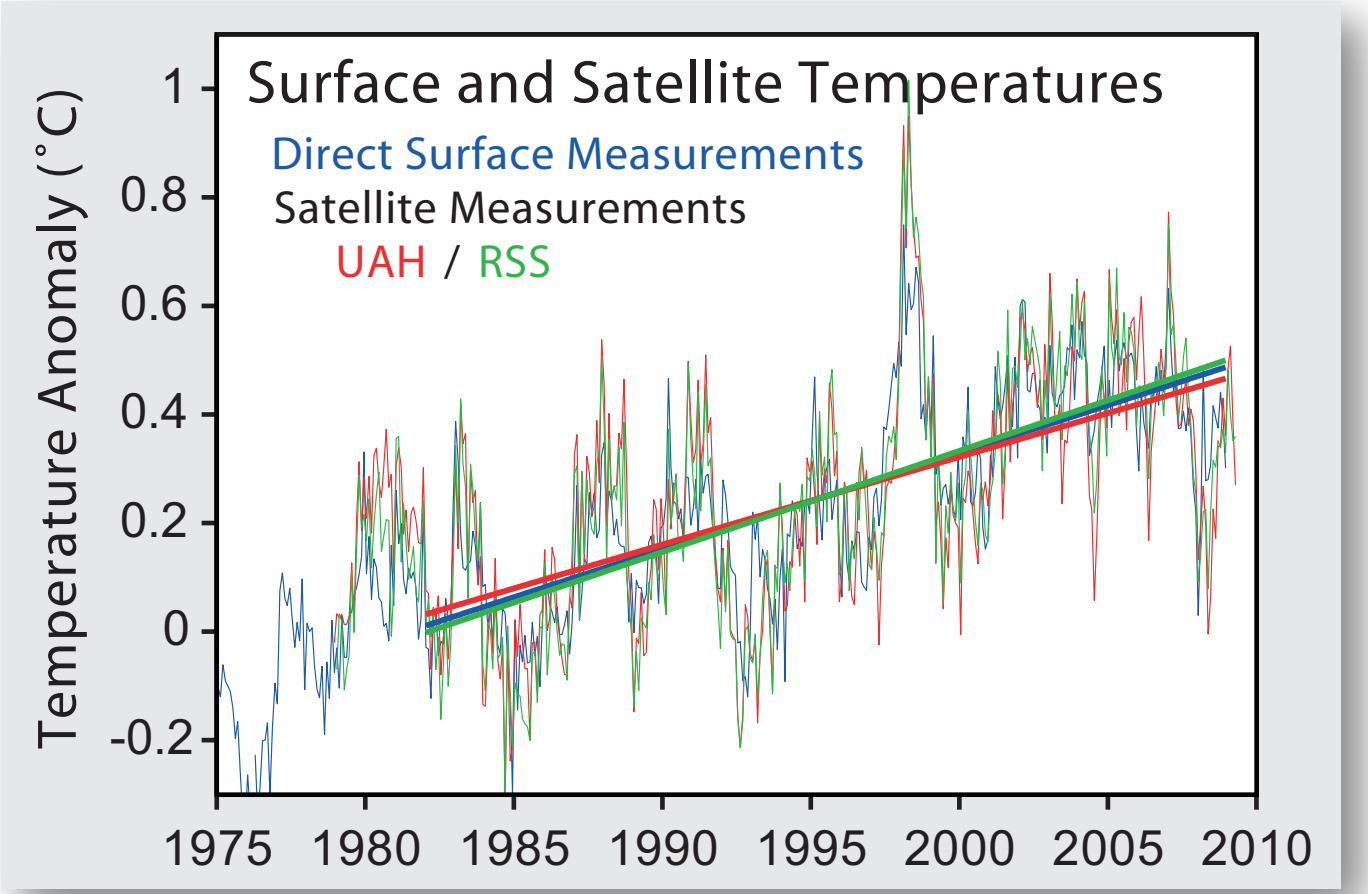
Source: W. Soon et al., "Environmental Effects of Increased Atmospheric Carbon Dioxide," Climate Research 13, no. 2 (1999): 153

Tilting the line: UAH D (1999) vs. UAH 5.2 (2005) and RSS 3.0 (2007)



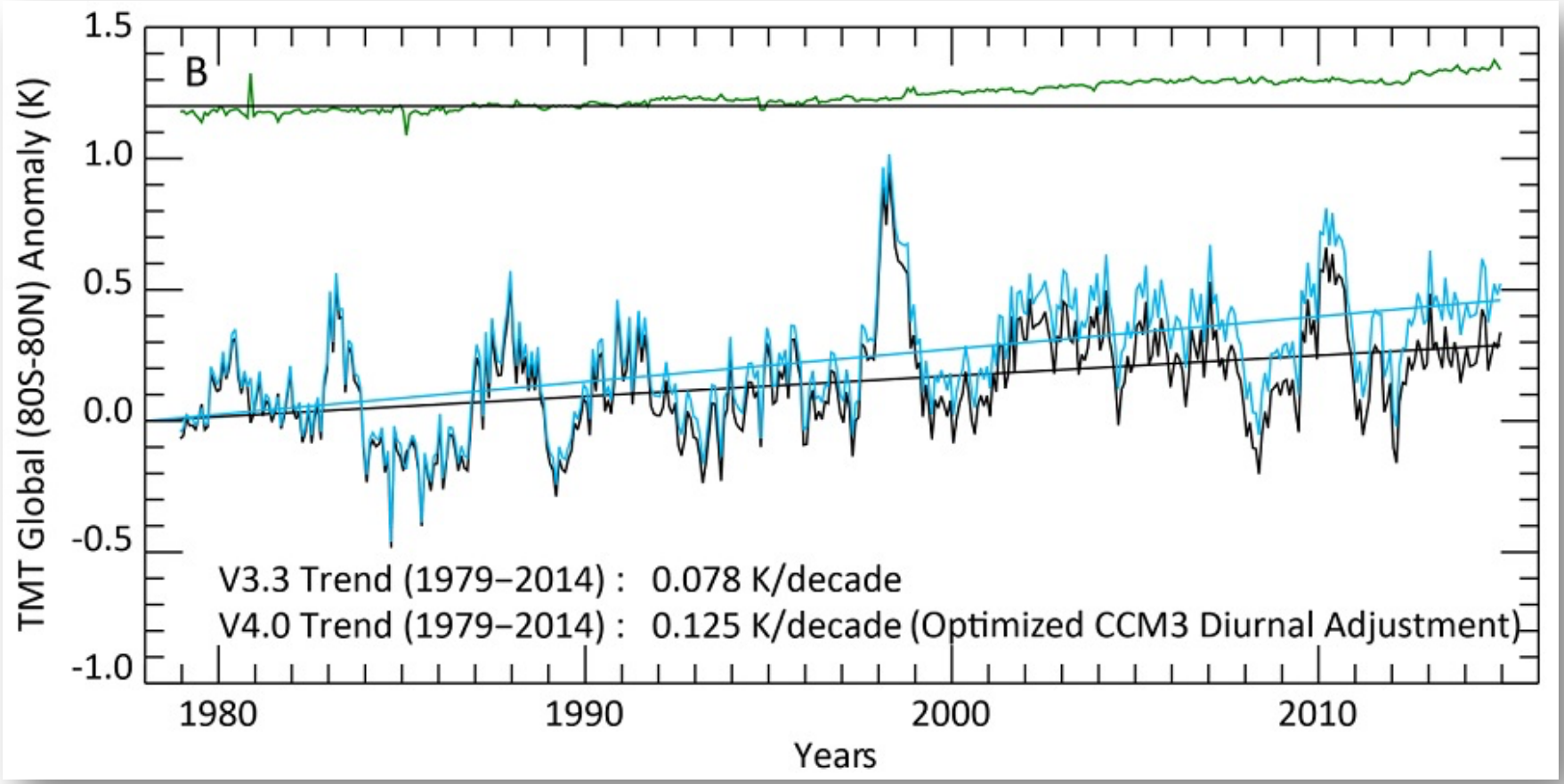
Source: Keller 2008, "Global Climate Change: A Review of this Mostly Settled Issue," Stochastic Environmental Research and Risk Assessment 23:5, 643-676

Tilting the line: surface vs. satellite TLT (lower troposphere), 2008



Global temperature anomaly in surface data (HadCRUT3) vs. lower-troposphere MSU data from UAH V5.2 and RSS V3.2. Graphic by Robert A. Rohde.

TMT (middle troposphere): RSS v4.0 (blue, applying new diurnal adjustment) vs. RSS 3.3



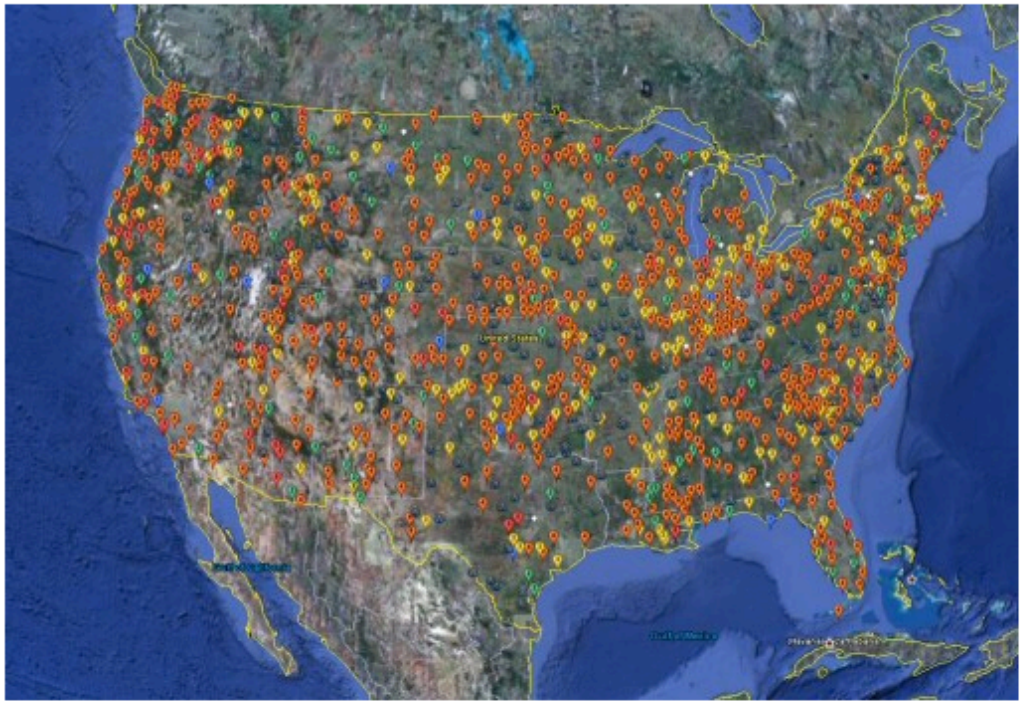
Source: Mears & Wentz (2016), "Sensitivity of Satellite-Derived Tropospheric Temperature Trends to the Diurnal Cycle Adjustment," *Journal of Climate*

surfacestations.org

A resource for climate station records and surveys

NEWS Updated 07/16/2009

Surfacestations project reaches 82% of the network surveyed. 1003 of 1221 stations have been examined in the USHCN network. The Google Earth map below shows current coverage.



CRN Rating key	1	2	3	4	5	
Estimated Error in °C (per NOAA)	Error ≤1°C	Error ≤1°C	Error ≥1°C	Error ≥2°C	Error ≥5°C	Unrated
Quality	Best	Good	Fair	Poor	Worst	Closed

Reference for site ratings: NOAA's Climate Reference Network Site Handbook Section 2.2.1

HOME CONTACT

Get Involved! help us document weather stations in the USA and the world.

Odd and irregular observing Sites looking at some of these observing sites you have to wonder: "what were they thinking"?

Resources links to useful and pertinent documents, images, drawings, specifications, and web sites.

Visit the blog to see highlighted examples of poorly sited stations in the "How Not to measure Temperature" series.

Site launched on 06/04/07

Progress as of 07/16/2009

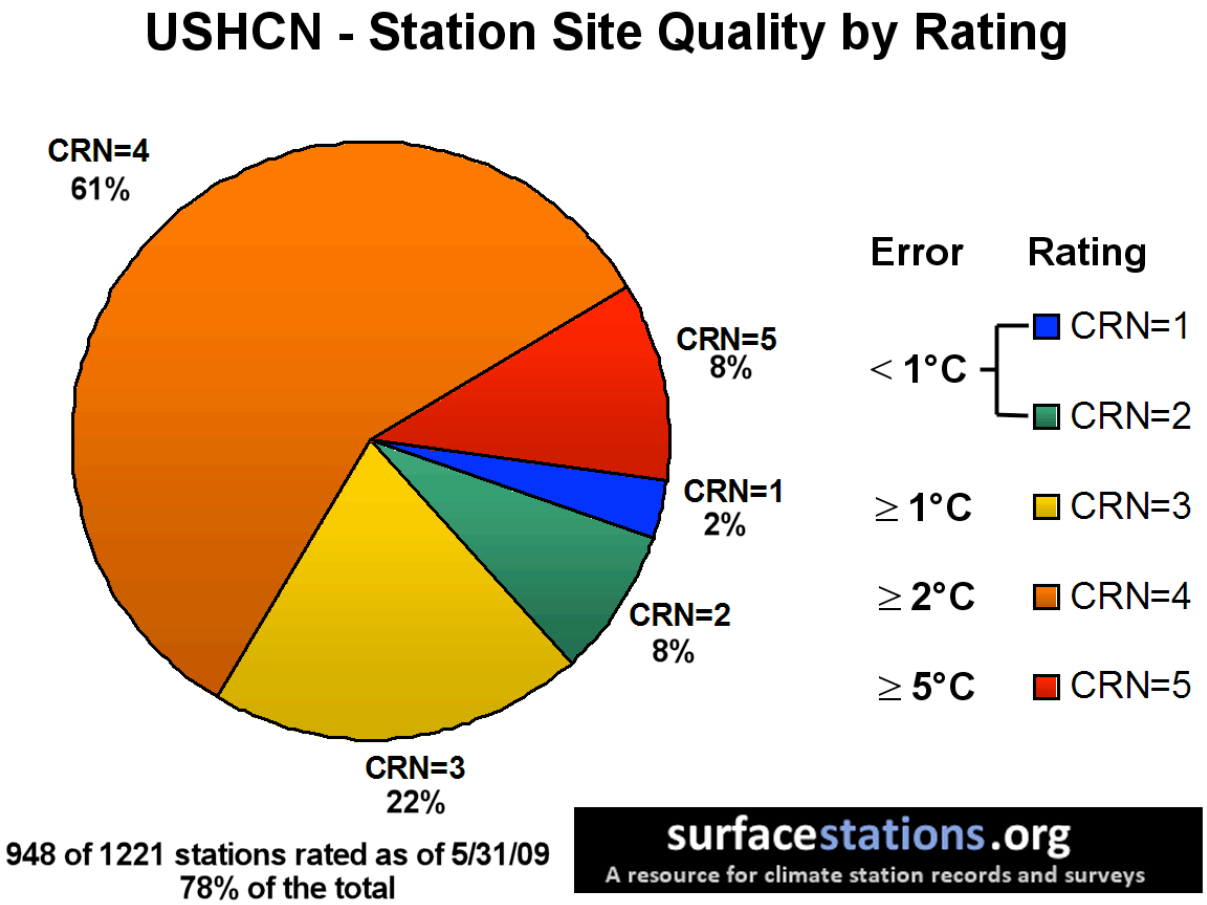
USHCN Sites surveyed so far:

1003

From A. Watts, “Is the U.S. Temperature Record Reliable?”, Heartland Institute, 2009

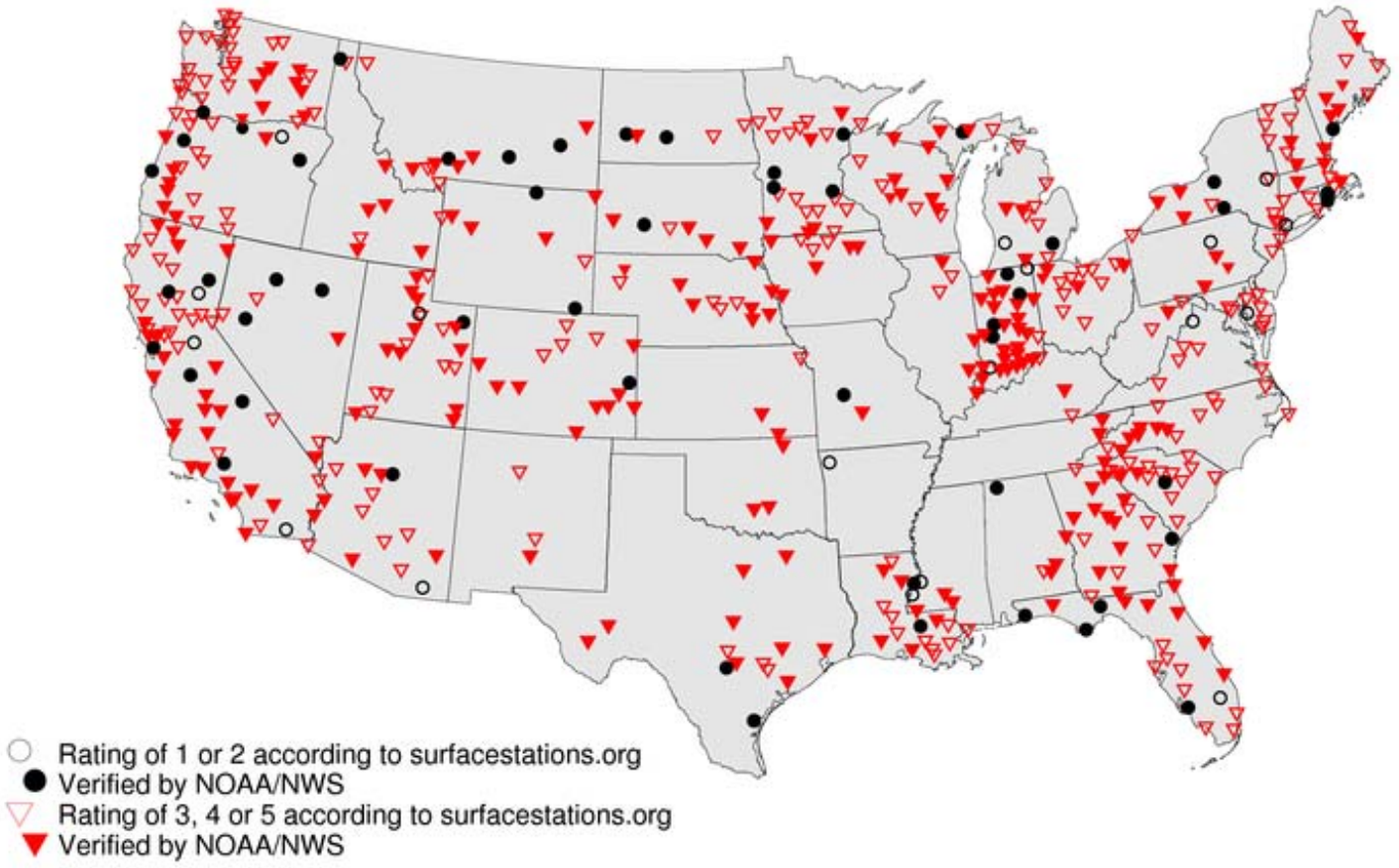
MMTS = Maximum/Minimum Temperature System (electronic thermistor)

“We were shocked by what we found… 9 of every 10 stations are likely reporting higher or rising temperatures because they are badly sited (p. 3)”



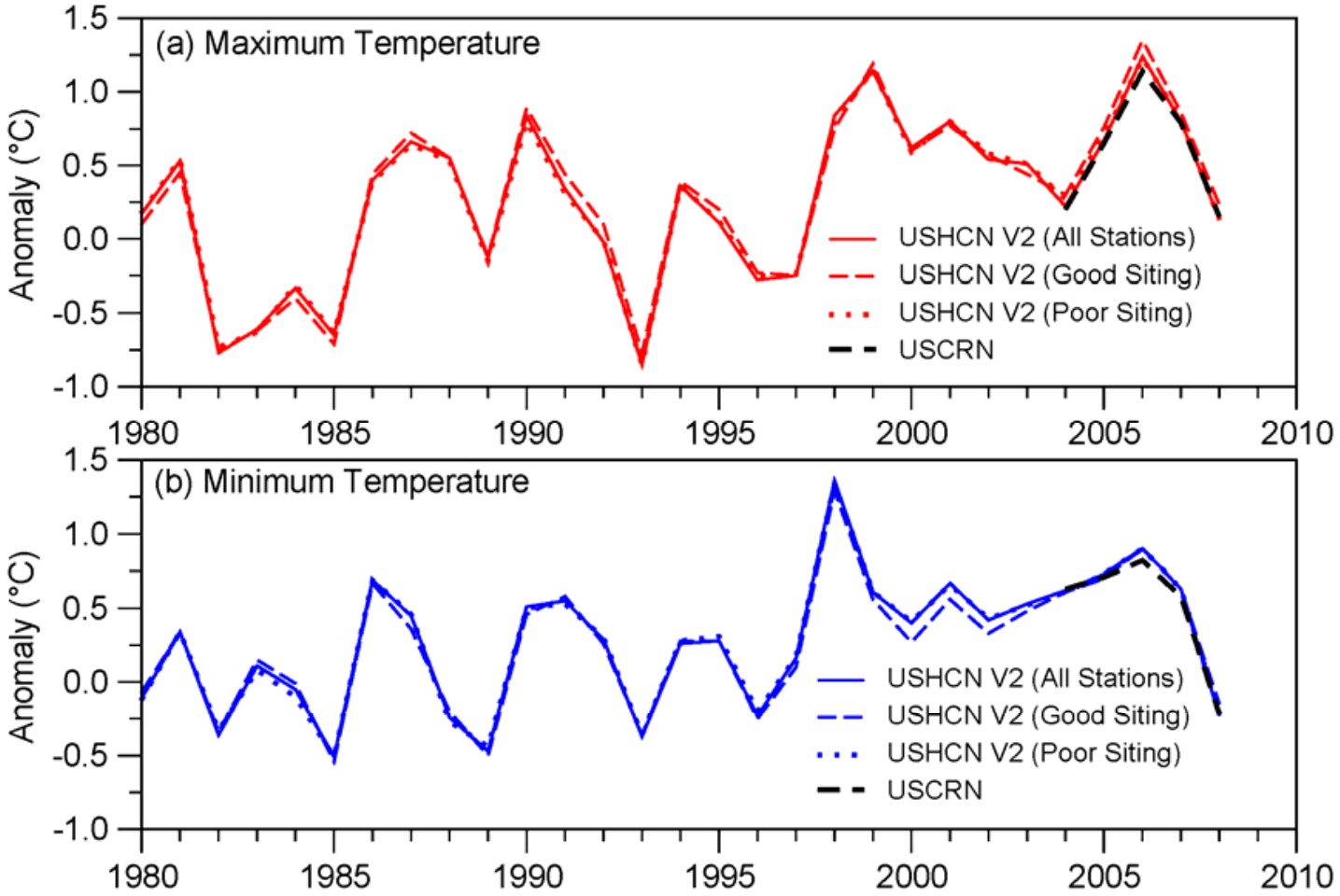
"Station quality ratings obtained from NOAA/NCDC via this source: Climate Reference Network Rating Guide - adopted [sic] from NCDC Climate Reference Network Handbook, 2002, specifications for siting (section 2.2.1)"

Menne et al. (2010), Fig. 1



USHCN exposure classifications according to surfacestations.org (circles and triangles). Filled symbols are in agreement with independent assessments by NOAA/National Weather Service Forecast Office personnel. …Ratings 1 and 2 are treated as "good" exposure sites; ratings 3, 4 and 5 are considered "poor" exposure.

Source: "V1.05 USHCN Master Station List". (Downloaded from www.surfacestations.org in June 2009. A more complete set of USHCN station classifications as referenced in Watts [2009] was not available for general use at the time of this analysis).



"Comparison of the [continental US] average annual (a) maximum and (b) minimum temperatures calculated using USHCN version 2 adjusted temperatures [Menne et al. 2009] and USCRN departures from the 1971-2000 normal. Good and poor site ratings are based on surfacestations.org."

Source: Menne et al., "On the reliability of the U.S. Surface Temperature Record," J. Geophys. Research (2010), Fig. 7

Menne et al. (2010)

- Conclusion: widespread poor site exposure in USHCN is real, *but*…

  "The bias in unadjusted maximum temperature data from poor exposure sites relative to good exposure sites is, on average, *negative*...

  Adjustments applied to USHCN Version 2 data largely account for the impact of instrument and siting changes, although a small overall residual negative ("cool") bias appears to remain…

  We find no evidence that … US temperature trends are inflated due to poor station siting."

Introduction

Weather forecasting

Computers change everything

Climate models and climate data

Data friction

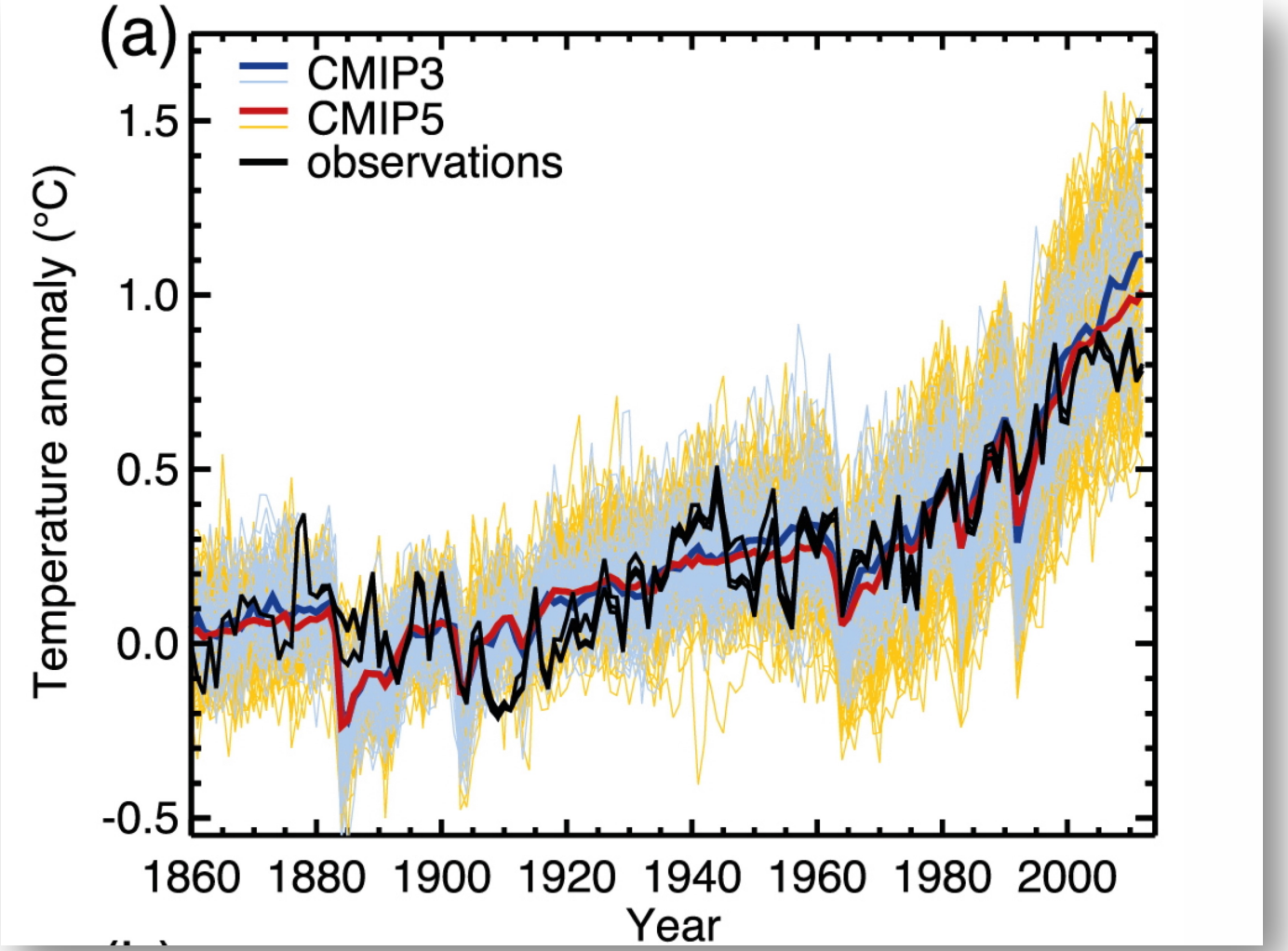
Climate controversies

Conclusion

Some lessons

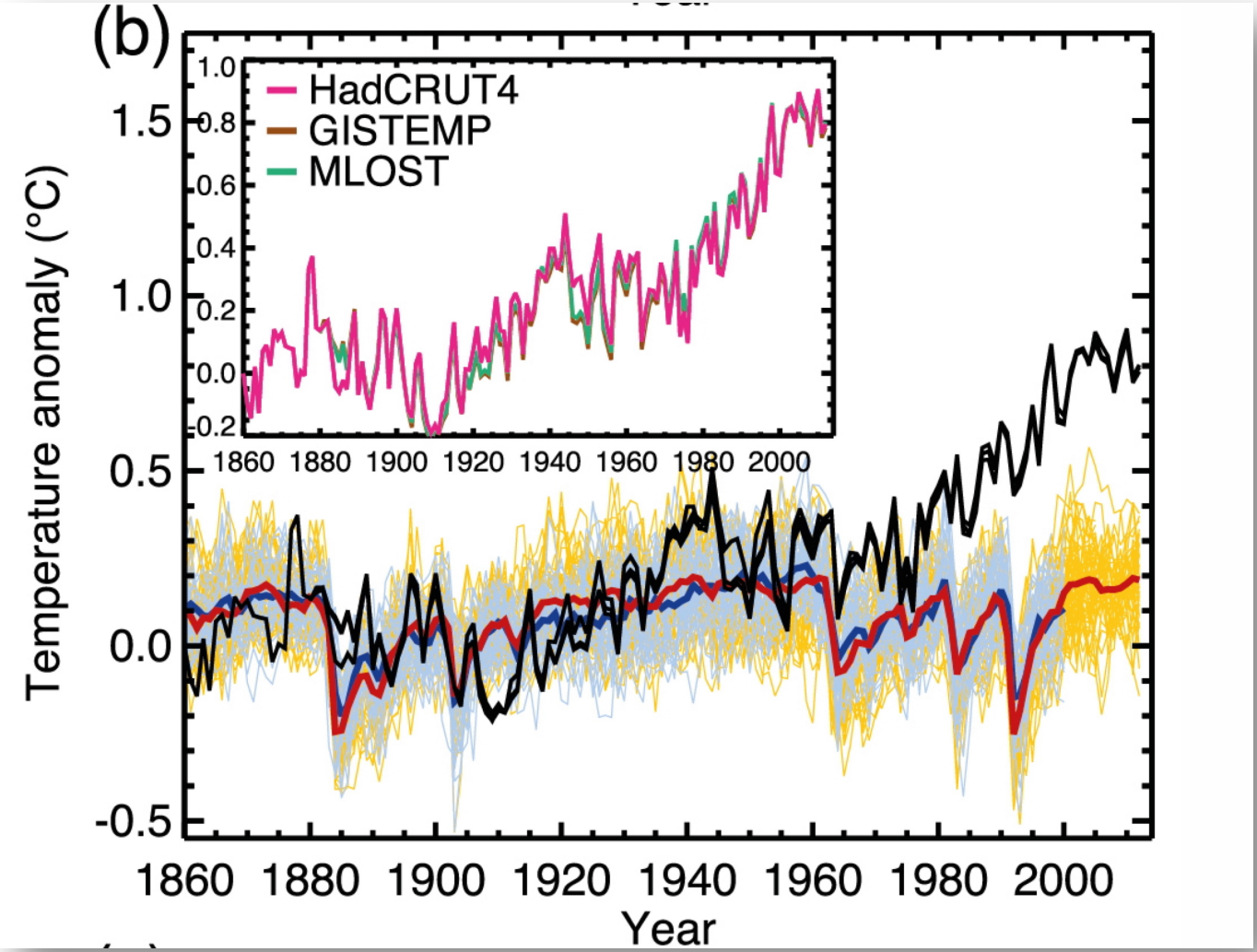
- Climate knowledge is real and well settled
  - Strong, stable, very old infrastructure
  - *Not* "uncertain" in ordinary sense
- Data are made, not born
  - Global data are *always* produced by a combination of observation and modeling
- Infrastructural inversion is a fundamental method
  - Constantly surfaces data problems and generates new versions of data
- Despite convergence, data instability can always be used to provoke doubt

Climate models vs. observations



Source: IPCC AR5 (2013), Figure 10-1

Climate models vs. control Earth



Source: IPCC AR5 (2013), Figure 10-1

Versions of the future